Universidade de Aveiro
2018
Departamento de
Electrónica, Telecomunicações e Informática

Carlos Alexandre da Costa Neves

Privacidade do utilizador em sítios na web dos municípios portugueses

Carlos Alexandre da Costa Neves

Privacidade do utilizador em sítios na web dos municípios portugueses

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia de Computadores e Telemática, realizada sob a orientação científica do Professor Doutor Hélder Gomes e do Professor Doutor Fábio Marques.

o júri / the jury

presidente / president	**Prof. Dr. André Zúquete** Professor Auxiliar, Universidade de Aveiro
vogais / examiners committee	**Prof. Dr. Telmo da Silva** Professor Auxiliar, Universidade de Aveiro
	Prof. Dr. Fábio Marques Professor Adjunto, Universidade de Aveiro

agradecimentos

Quero agradecer a todos os que me acompanharam nestes últimos meses, incluindo os meus pais, a minha namorada e o meu primo. Não teria conseguido sem todos vocês.
Obrigado.

Palavras-Chave Privacidade, *internet*, pseudonimização, RGPD, *cookie*, *tracking*, legislação, cidadão, consentimento, *crawler*, automatização.

Resumo

Cada vez mais os sítios na *web* são ferramentas usadas para a interação das mais variadas instituições com os cidadãos. Nestas interações é habitual a recolha de informação dos cidadãos, que pode ser feita com ou sem o conhecimento destes e podendo envolver dados pessoais. Com o objetivo de proteger a privacidade dos utilizadores, tem sido desenvolvido um conjunto de legislação a que as comunicações através da *Internet* estão sujeitas, de que são exemplos a lei de proteção de dados pessoais e a lei de proteção dos dados pessoais nas comunicações eletrónicas, e, para determinadas áreas, regras de boas práticas que os sítios na *web* devem respeitar.

Esta dissertação tem como objetivo o estudo da conformidade de sítios na *web* de uma determinada área homogénea (por exemplo, os *sites* dos municípios, das universidades, etc.) com a legislação e boas práticas existentes, no que se refere à privacidade na interação com os utilizadores. Nomeadamente pretende-se que a verificação da conformidade possa ser feita através da utilização de ferramentas, a desenvolver, que automatizem o processo de verificação. Como exemplo, algumas características de possível observação são: a utilização de ligações seguras, a existência de *trackers*, a existência de políticas de privacidade e nível de consentimento.

Surgiu a necessidade de estudar várias ferramentas e respetivas características a fim de perceber se alguma se adequa à análise pretendida. Optou-se pela utilização do OpenWPM, com a necessidade de desenvolvimento de funcionalidades, para a recolha e consequente análise da informação necessária para averiguar a preocupação com a privacidade do utilizador na navegação em sítios *web* dos municípios portugueses.

Concluiu-se que a maioria dos municípios recolhem informação sobre os seus utilizadores sem que sejam respeitadas as normas, nomeadamente no que toca à necessidade informação dos utilizadores ou da requisição de consentimento.

keywords Privacy, *internet*, pseudonymization, GDPR, cookie, tracking, legislation, citizen, consent, crawler, automation.

Abstract

Increasingly, websites are becaming tools used for the interaction of the most varied institutions with citizens. In these interactions institutions may gather information from citizens, which can be done with or without the knowledge of these and may involve personal data or not. In order to protect the privacy of users, a set of legislation has been developed to which *Internet* communications are subject, such as the law of protection of personal data and the law of protection of personal data in electronic communications and, for certain areas, rules of good practice that Websites must respect.
This dissertation aims to study the conformity of Websites in a given homogeneous area (eg, municipalities, universities, etc.) with existing legislation and good practices regarding privacy in the interaction with users. In particular, it is intended that the verification of conformity can be done through the use of tools to automate the verification process. As an example, some characteristics of possible observation are: the use of secure connections, the existence of trackers, the existence of privacy policies and consent level.
There was a need to study various tools and their characteristics in order to see if any fit the required goal. We chose to use OpenWPM, with the need to develop extra functionalities, for the collection and consequent analysis of the information necessary to ascertain the concern with the privacy of the user when navigating in websites of the portuguese municipalities.
It was concluded that most municipalities collect information about their users without complying with the rules, especially about informing the users or the request for consent.

Índice

Índice de Tabelas e Figuras

Tabelas

Figuras

Índice de Gráficos

Glossário

ACEPI	Associação de Economia Digital
AIPD	Avaliação de Impacto sobre a Proteção de Dados
AP	Administração Pública
API	Application Programming Interface ou Interface de Programação de Aplicação
CE	Comissão Europeia
CERN	Organização Europeia para a Pesquisa Nuclear conhecida como CERN (antiga sigla para *Conseil Européen pour la Recherche Nucléaire*)
DHCP	*Dynamic Host Configuration Protocol*
DNT	*Do Not Track*
DOM	Modelo de Objeto de Documento ou *Document Object Model*
DPA	*Data Protection Analyst*
DPO	*Data Protection Officer*
EUA	Estados Unidos da América
HTTP	*Hypertext Transference Protocol* ou Protocolo de Transferência de Hipertexto
HTTPS	HTTP *Secure* ou Protocolo de Transferência Segura de Hipertexto
IDC	International Data Corporation
IP	*Internet Protocol*
LSO	*Locally Shared Objects*
NAT	*Network Address Translation*
PDF	*Portable Document File*

RGPD/ GDPR	Regulamento Geral Proteção Dados ou, em inglês, *General Data Protection Regulation*
SSL	*Secure Sockets Layer*
TCP	*Transmission Control Protocol*
TLS	Transport Layer Security
EU	União Europeia
URL	Localizador Uniforme de Recursos, em inglês, Uniform Resource Locator
WPM	Medição de Privacidade na *web*, em inglês, *Web privacy measurement*

1 Introdução

Desde que surgiu, nos anos 60, muito mudou na *Internet* até aos dias de hoje. O contributo de Tim Berners-Lee com a criação do primeiro *website*[1], em agosto de 1991, aliado à criação da *World Wide Web* [1], com a Organização Europeia para a Pesquisa Nuclear (CERN), foram passos cruciais para possibilitar o desenvolvimento das ferramentas e tecnologias disponíveis atualmente. A *internet* revolucionou, globalmente, a vida quotidiana de muitas empresas, instituições e indivíduos tornando-se um importante elo de ligação entre todos estes intervenientes.

Num mundo globalizado e conectado, a capacidade de pesquisa e partilha de informação constante e atualizada, representa uma mudança nos comportamentos e dinâmicas sociais, possuindo um alcance e abrangência ímpares. O crescente número de utilizadores da *web* implica um aumento exponencial de transações de informação. Uma transação na *web* é qualquer processo que induz uma transferência de informações entre duas ou mais entidades; algumas destas informações podem ser, eventualmente, de caráter pessoal tornando-se indispensável a existência de medidas regulatórias que visem a proteção da identidade e privacidade do utilizador.

Um dos maiores desafios deste século é a aplicação de normas para a proteção dos dados dos utilizadores da *web*, dada a sua natureza inerentemente aberta, não determinística e a sua complexidade.

[1] http://first-website.web.cern.ch/blog/first-url-active-once-more (acedido em 18/01/2018)

A *internet*, enquanto tecnologia, aliada ao elevado número de transações que ocorrem no ciberespaço, representa por si só uma barreira à monitorização e determinação da conformidade com as normas.

1.1 Contextualização

A constante evolução da tecnologia e o seu entrosamento no quotidiano dos cidadãos em praticamente todo o Mundo, torna difícil definir conceitos como privacidade ou a sua invasão [2]. Aceite como exigência direta de cada indivíduo, a privacidade é, cada vez mais, consagrada como um direito fundamental presente em documentos de natureza jurídica [3].

Na sua origem etimológica, privacidade refere-se ao que é privado ou íntimo. É o direito à reserva de todo o tipo de dados pessoais, podendo entender-se também como direito a controlar a exposição de informações acerca de si próprio.

Em Portugal, segundo a Constituição da República [4], o direito à privacidade é a possibilidade de assegurar ao indivíduo a reserva de um espaço da sua vida privada e familiar em que esteja a salvo da intromissão de terceiros. A extensão deste espaço e a sua preservação variam de acordo com o posicionamento social.

No seu artigo 26º, a Constituição consagra direitos pessoais fundamentais, descrevendo-os no seu ponto 1. Nos pontos seguintes são dadas garantias contra obtenção e utilização abusiva dos dados privados, dando especial atenção à identidade genética do ser humano na criação ou utilização de tecnologias ou experimentação científica.

No âmbito da utilização informática de dados, no artigo 35º da Constituição está ressalvado o direito de acesso aos dados pessoais informatizados para consulta, retificação ou atualização bem como o direito de conhecer o objetivo da recolha e tratamento desses dados. Nos segundo e terceiro pontos deste artigo estabelece-se que a lei define o conceito de dados pessoais e as condições aplicáveis ao seu tratamento automatizado, conexão, transmissão e utilização, limitando a utilização discriminatória – com base em convicções filosóficas ou políticas, filiação partidária ou sindical, fé religiosa, vida privada e origem étnica – garantindo a sua proteção, nomeadamente

através da Comissão Nacional de Proteção de Dados (CNPD[2]), enquanto entidade administrativa independente.

Está ainda previsto neste artigo a impossibilidade de acesso a dados pessoais de terceiros, salvo em casos excecionais. No sétimo e último ponto do artigo 35º, é ainda definido que os ficheiros manuais de registos de dados têm a mesma condição legal dos informatizados ou automatizados.

No que diz respeito à violação dos direitos à privacidade, podemos classificá-la em diferentes níveis: informação do espaço pessoal e territorial do corpo; o campo psicológico e moral; a dignidade e respeito de cada indivíduo.

Pode resumir-se a definição de dados pessoais, a todos e quaisquer dados relativos a pessoas singulares identificadas ou identificáveis, no entanto o termo "dados pessoais" é muito mais amplo, podendo dizer respeito a nome, morada, idade, estado civil, *email*, situação patrimonial ou financeira, entre outros. Todas as formas de registo desses dados estão contempladas, seja em suporte papel, som, imagem ou suporte informático [5].

Para se poder entender os riscos da quebra de privacidade de dados, é necessário que sejam previamente identificados e caracterizados os diferentes tipos de dados relativos a cada indivíduo.

De forma geral, todas as entidades, sejam singulares ou coletivas possuem dados gerais de identificação, tais como:

- nome;
- morada;
- Número de Identificação Civil;
- Número de Identificação de Segurança Social;
- Número de Identificação Fiscal;
- contacto telefónico;
- *email*.

Existem, também, dados pessoais específicos de determinadas organizações como bancos, seguradoras e outras atividades que registam dados como:

[2] http://cnpd.pt (acedido em 20/1/2018)

- número de cliente;
- IBAN;
- número de sócio;
- número mecanográfico.

Outro tipo de dados pessoais são os dados biométricos que se dividem em duas categorias - físicas e comportamentais. Na primeira categoria incluem-se informações e características únicas de cada ser humano:

- impressões digitais;
- geometria de mãos e dedos;
- veias;
- face;
- íris;
- retina;
- voz (timbre vocal).

Apesar de não serem eventualmente tão evidentes, os dados biométricos comportamentais são informações únicas de cada ser. São bons exemplos destas características:

- a assinatura;
- a escrita manual;
- a escrita a computador;
- voz (volume e tom);
- gestos corporais.

Hoje, a utilização de dados biométricos é comum em diversas organizações permitindo identificar e autenticar os indivíduos em sistemas de registo de assiduidade ou em controlo de acesso a locais restritos.

Estando o Ácido Desoxirribonucleico (ADN) considerado como dado genético, a identificação por esta via não é considerada uma tecnologia biométrica de reconhecimento por ser ainda um processo demorado na sua análise - apesar de ser uma análise de ampla origem uma vez que as moléculas de ADN estão presentes em todas as células do corpo. Esta informação genética é exclusiva de cada indivíduo. Por poder ser traduzida numa sequência numérica, a informação de ADN - os perfis de ADN - pode ser armazenada numa base de dados informatizada. Existem ainda dados

pessoais de origem clínica, que são aqueles que dizem respeito às informações necessárias para a execução de atos médicos.

Hoje em dia, a capacidade de controlo da divulgação dos nossos dados é cada vez menor, devido à ampla utilização de plataformas *online*. Esta lacuna na capacidade de controlo deve-se, muitas vezes, à falta de noção na distinção entre mundo real e mundo virtual; distinção essa que aumenta a frequência com que é desprezada a privacidade e a proteção de dados, tanto no desenvolvimento aplicacional, como na sua utilização.

As empresas estão cada vez mais presentes no mundo digital, disponibilizando aos seus clientes cada vez mais serviços *online*. Aderir a estes serviços, efetuar compras *online* ou subscrever *newsletters* de *blogs* ou outros *sites* de interesse são atividades comuns no nosso dia-a-dia.

Para os serviços públicos esta mudança de paradigma *offline vs online* também é uma realidade: é cada vez mais comum a utilização da *internet* em substituição da deslocação aos espaços físicos das entidades governamentais. É agora possível, praticamente em qualquer lugar e/ou equipamento, fazer pedidos de certidões e declarações de caráter legal ou mesmo solicitar procedimentos administrativos (como processos de casamento e divórcio, renovação do cartão de cidadão, carta de condução, entre outros).

A Associação de Economia Digital (ACEPI) e o International Data Corporation (IDC), no Estudo Anual da Economia e da Sociedade Digital em Portugal [6], analisam a evolução da utilização da *internet*, a presença *online* das organizações, o impacto da mobilidade na economia digital e o comércio eletrónico. Comparando este estudo com o "Estudo sobre Local e-Government em Portugal" da Universidade do Minho [7], é possível perceber que, percentualmente, em todos os itens analisados, a Administração Pública (AP) totaliza menos pontos que as empresas comerciais – em 2016, como demonstra a Figura 1 [7], apenas 89% dos organismos da AP possuía *site* próprio enquanto as empresas somam 92%.

Edição	Ano	Número Total de Câmaras Municipais	Sítios Web Identificados	% Câmaras com Sítio Web	Referência do Estudo
1.ª	1999	305	153	50%	(Santos e Amaral 2000)
2.ª	2001	308	222	72%	(Santos e Amaral 2003)
3.ª	2003	308	259	84%	(Santos e Amaral 2005)
4.ª	2005	308	303	98%	(Santos e Amaral 2006)
5.ª	2007	308	306	99%	(Santos e Amaral 2008)
6.ª	2009	308	308	100%	(Santos e Amaral 2012)
7.ª	2012	308	308	100%	(Soares et al. 2014b)
8.ª	2014	308	308	100%	(Soares et al. 2016)
9.ª	2016	308	308	100%	*presente relatório*

Figura 1 – Edições do estudo "Presença na Internet das Câmaras Municipais Portuguesas" [7]

Analisando especificamente a utilização de *websites* por parte das câmaras municipais [8], percebemos que em 1999, no total dos 305 municípios (aumentou em 2001 para 308 câmaras) metade possuía já um sítio na *web*; o estudo, feito de dois em dois anos, revela que a percentagem de câmaras com *website* aumentou progressivamente até 2009, atingindo os 100%.

Com base nos dois estudos citados anteriormente, podemos concluir que atualmente todas as câmaras possuem já um *website* próprio. É através desta ferramenta que cada câmara municipal informa os seus cidadãos das decisões tomadas em órgãos deliberativos, nomeadamente na publicação de editais, atas, documentos relativos a taxas e licenças e dá a conhecer outras informações oficiais de origem fiscal e de gestão. Na maioria dos casos, os municípios permitem aos utilizadores submeter e consultar pedidos à sua câmara municipal. Endereçados aos diversos setores da atividade dos municípios, os pedidos são submetidos de diversas formas - seja pelo preenchimento de modelos em PDF devolvidos via *site* ou por *email*, seja em formulários próprios dentro das respetivas áreas de utilizador.

1.2 Motivação

Atentando à falta de preocupação no processamento de dados no desenvolvimento de aplicações e de forma a salvaguardar a privacidade dos cidadãos, é necessário compreender e verificar o cumprimento das diretivas a que as entidades estão sujeitas. Em Portugal, a Lei nº 67/98, da Proteção de Dados Pessoais [9] e a Lei nº 46/12, da Proteção dos Dados Pessoais nas Comunicações Eletrónicas [10], colocam

na esfera jurídica as diretivas europeias sobre a proteção das pessoas singulares, no que diz respeito ao tratamento e circulação desses dados.

No entanto, convém notar-se que estas normas legais, em Portugal e no resto da Europa, não têm como objetivo criar um quadro jurídico que possa abordar o futuro tratamento de dados e os desafios da privacidade. Em 2010, a Comunidade Europeia (CE) define a estratégia para a revisão da legislação sobre a proteção de dados e publica um comunicado [11] onde mostra indícios de que a União Europeia (UE) necessita de novas regras para enfrentar a mudança do mundo globalizado e permanecer relevante para as tecnologias inovadoras [12]. Em 2012, é publicada a primeira proposta do Regulamento Geral de Proteção de Dados (RGPD) que inclui a maioria das regras que entraram em vigor em maio de 2018 [13]. Não obstante, muitas das regras deste primeiro projeto, sofreram alterações por serem consideradas excessivamente restritas ou, por outro lado, muito desvinculadas. Até à sua votação e aprovação em abril de 2016 [14], foram muitas as discussões e alterações a este regulamento.

O RGPD foi concebido para modernizar e harmonizar a legislação da UE em matéria de proteção de dados. O desenvolvimento da tecnologia e o ambiente inovador exigiram mudanças significativas em relação à privacidade dos dados processados. Do ponto de vista das empresas e organizações, o RGPD significa sem dúvida, de um lado, novas obrigações e mudanças estruturais para a conformidade, mas, por outro, o mais importante, certeza e segurança; eliminação de obstáculos e encargos relativos à transferência de dados [15]. Ao adotar o Regulamento – que entrou em vigor automaticamente em todos os países da UE, sem que seja necessária a transposição para as leis nacionais –, todas as instituições da UE reconhecem a inovação como estando presente em todos os aspetos do sistema empresarial e colocando as bases da legislação moderna em harmonia com o mundo tecnologicamente avançado.

1.3 Problema

Com a entrada em vigor do RGPD, a importância na identificação de boas práticas nas interações *web* tornou-se ainda mais relevante, na medida em que todas as instituições europeias com presença *online* têm agora um conjunto de normas a

seguir, tornando-se assim, necessário identificar mecanismos que garantam a conformidade com a lei.

O impacto do RGPD tanto nas organizações como nas pessoas singulares é bastante significativo e complexo, produzindo novas obrigações para as entidades que tratam dados pessoais e estabelecendo um novo paradigma no recolha e tratamento desses mesmos dados.

Representando não só o poder local (os cidadãos) mas também sendo, em cada município, um órgão de gestão governamental, as câmaras municipais devem ser e dar exemplo da aplicação das regras e boas práticas definidas no RGPD, não só, mas também nos seus sítios *web*.

Uma vez que o Regulamento reforça os direitos dos cidadãos, cria novos procedimentos e novas obrigações para todas as entidades da AP, é necessário conhecer o seu impacto concreto e específico no âmbito dos órgãos da AP local, sobretudo nas câmaras municipais.

1.4 Objetivo

Este trabalho tem como objetivo, propor um modelo de controlo da interação na *web*, baseando-se nos mecanismos de caracterização e segmentação dos utilizadores; na frequência de utilização desses mecanismos; na discrepância entre os mecanismos utilizados e aqueles que são divulgados ao utilizador e, por último, mas tão ou mais importante, a requisição prévia da aceitação explícita da utilização desses mecanismos.

O modelo pretende automatizar o processo de verificação destas características, simplificando o processo de análise, reduzindo-o apenas à avaliação de informação útil. Pretende-se que o modelo seja aplicado especificamente a uma área homogénea – organizações com interesses comuns -, com o intuito de obter uma visão geral sobre o seu nível de conformidade com as normas europeias, garantindo assim uma análise concreta e focada, numa avaliação justa entre pares.

1.5 Contribuição

Considera-se que a presente dissertação e o trabalho desenvolvido são, acima de tudo, um contributo para a consciencialização para a necessidade de preocupação com a privacidade dos dados, não só na utilização, mas também no desenvolvimento de sítios *web*.

O estudo das normas europeias em vigor, nomeadamente através da investigação e análise ao RGPD (na secção 2.2), e o estudo dos mecanismos de *tracking* e de defesa (nas secções 2.4 e 2.6, respetivamente), permite aos interessados no tema, um adquiri e aprofundar de conhecimentos sobre as normas e procedimentos para que os *websites* sigam essas normas.

O foco nos municípios portugueses permite aos responsáveis desta área homogénea percecionar o estado da conformidade dos seus sítios na *web.* Através da interpretação da informação recolhida é possível determinar quais os pontos em que os municípios devem alterar a sua conduta a fim de seguir as normas para a proteção dos dados dos seus utilizadores.

Espera-se que esta dissertação e os dados recolhidos sejam incentivo e base de trabalho para o desenvolvimento de outros estudos no futuro, quer ao nível da continuidade da avaliação da privacidade do utilizador nos *websites* dos municípios, quer ao nível da avaliação da presença *web* de outras áreas.

1.6 Estrutura da Dissertação

Este documento está organizado em 6 capítulos, sendo o primeiro e o último a introdução e a conclusão, respetivamente.

O segundo capítulo apresenta o estado da arte que aborda seis tópicos principais: privacidade no tratamento de dados, o RGPD, comunicações seguras, *tracking* na *web*, medição de privacidade na *web* e mecanismos de defesa contra *tracking* na *web.*

O terceiro capítulo propõe uma solução, identificando claramente, nas primeiras 5 secções principais, os pontos a serem avaliados para garantir a conformidade com as normas. Na última secção são comparadas as principais soluções a considerar e é definida a plataforma a utilizar.

O quarto capítulo descreve, por temáticas, as implementações necessárias à ferramenta escolhida para garantir a automatização da recolha dos dados necessários à análise da privacidade do utilizador em sítios na *web* dos municípios portugueses.

No quinto capítulo são descritos e analisados os resultados recolhidos na navegação nos sítios *web* dos municípios portugueses.

2 Estado da Arte

As informações que podem ser obtidas nas interações na *web* são de natureza variada, bem como os mecanismos para a sua obtenção [16]. Na análise à privacidade do utilizador em sítios na *web* dos municípios, importa definir previamente dois momentos no controlo de dados – a recolha de dados e o tratamento ou processamento de dados. Após esta definição serão abordadas a privacidade dos dados e a sua proteção na legislação nacional e europeia; os pontos-chave do RGPD; os mecanismos de *web tracking* e os mecanismos de defesa.

Considera-se recolha de dados qualquer tipo de aquisição ou arquivo de dados, que poderá enquadrar-se numa forma lícita quando o consentimento é específico, objetivo e explícito.

O tratamento de dados é a manipulação ou aglomeração dos dados recolhidos para obter informações relativas ao titular de dados. Para que seja possível esta operação, o consentimento do respetivo titular deve ter ocorrido inequivocamente antes do processamento ocorrer.

2.1 Privacidade no Tratamento de Dados

Para ser compreensível a análise do nível de intrusão a dados pessoais de cidadão é necessário identificar certos conceitos e mecanismos. Serão abordados temas relativos a políticas de privacidade em vigor, a mecanismos de *web tracking* e comunicações seguras.

Hoje em dia, devido à quantidade e regularidade no pedido de dados pessoais, a privacidade é uma das maiores preocupações quer do ponto de vista dos utilizadores mas também das organizações [17]. Apesar disso, grande parte dos utilizadores da *internet* não tem noção da quantidade de informação que é recolhida nem o propósito da sua utilização. De forma a salvaguardar os utilizadores, foi aprovada legislação que visa aumentar o controlo do utilizador sobre o processamento dos seus dados, a redução dos dados solicitados e a responsabilização das entidades que procedam ao tratamento de dados.

Em 1995, a Diretiva 95/46/CE coloca na esfera jurídica a proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados. Tendo sido proposta já em 1990, esta diretiva refletia a necessidade de harmonizar a legislação dos diversos Estados-Membros, os quais, ou não tinham legislação nestas matérias, ou, quando existia, ofereciam graus de proteção variáveis ou falta de eficiência na aplicação da lei.

Em 2002, a Diretiva 2002/58/CE (Diretiva das Comunicações Eletrónicas) regulamenta "o tratamento de dados pessoais e a proteção da privacidade" no âmbito das "prestações de serviços de comunicações eletrónicas acessíveis em redes de comunicações públicas", i.e., esta diretiva aplica-se ao tratamento de dados pessoais feitos através da *internet*, enquanto serviço de comunicação eletrónica; não se aplicando, no entanto, se e quando os dados forem recolhidos e tratados em redes privadas – nestes casos é aplicada a Diretiva 95/46/CE.

Em 2009, a Diretiva 2009/136/CE altera um conjunto de diretivas, entre as quais a diretiva 2002/58/CE, definindo exceções à necessidade de consentimento do utilizador aquando da prestação de serviços requisitados pelo utilizador, quando seja

obrigatório o armazenamento de dados no terminal do utilizador para que esta prestação funcione corretamente.

Em 2012, em Portugal é revista a Lei nº 41/2004, relativa à Proteção de Dados Pessoais nas Comunicações Eletrónicas, passando a ser aplicável a Lei nº 46/2012, que transpõe a Diretiva 2009/136/CE e aplica à legislação nacional conceitos relativos à utilização da *internet* enquanto veículo de comunicação eletrónico. Esta Lei garante a segurança no processamento de dados, obrigando quem presta serviços de comunicações eletrónicas a aplicar "as medidas técnicas e organizacionais adequadas para garantir a segurança dos seus serviços, se necessário, no que respeita à segurança de rede" [10].

No mesmo ano, na Europa, a primeira proposta do RGPD já visava a clarificação da definição de consentimento explícito, a extensão da regulação a países fora da UE que regulam os cidadãos da UE e a imposição de coimas severas que poderão alcançar 4% das fontes de receita da entidade.

2.2 Regulamento Geral de Proteção de Dados

Revogando a Diretiva 95/46/CE e as suas transposições para as legislações nacionais, como é o caso da Lei 67/98 – Lei da Proteção de Dados Pessoais, o RGPD tem como principal objetivo proteger todos os cidadãos da UE das violações da privacidade num mundo cada vez mais orientado a dados. Pese embora alguns princípios fundamentais da privacidade de dados estejam considerados na diretiva anterior, o RGPD apresenta uma mudança significativa na definição de conceitos e na abrangência da aplicabilidade das normas europeias.

Daniel Solove esquematiza num *Whiteboard* [18] ilustrado na Figura 2, os diferentes pontos-chave do Regulamento, a sua aplicabilidade, os direitos dos utilizadores e os impactos sobre as organizações.

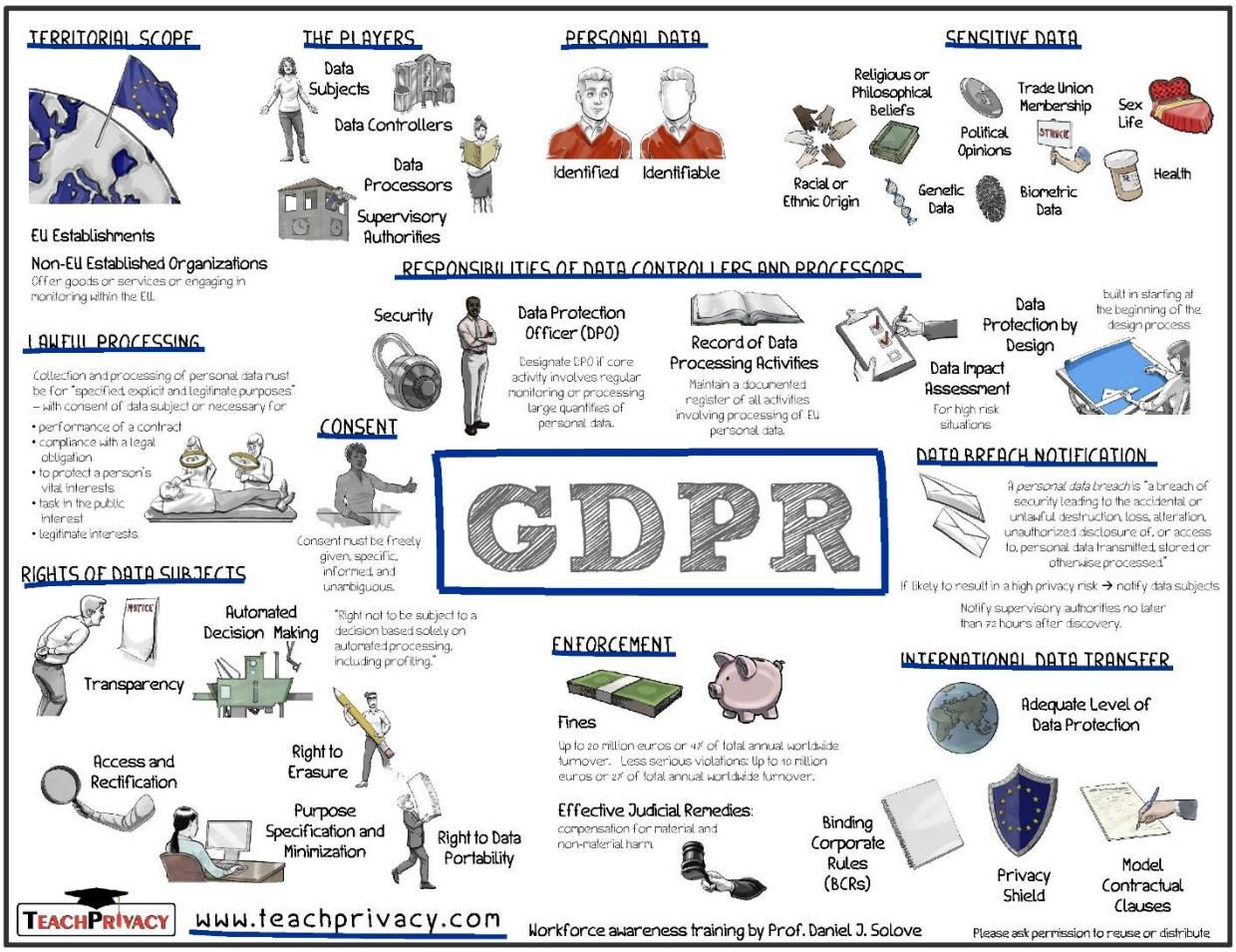


Figura 2 – Resumo visual das noções básicas do RGPD, direcionado a empresas e organizações [18]

2.2.1 Dados Pessoais

Colmatando lacunas na legislação anterior, este Regulamento passa a incluir na definição de dados pessoais, qualquer informação, de qualquer natureza e independentemente do seu suporte, mesmo que em som e imagem, relativa a um titular dos dados, i.e., uma pessoa singular identificada ou identificável. Considera-se identificada a pessoa associada a um identificador único e identificável uma pessoa que possa ser identificada indiretamente, através de elementos específicos da sua identidade física, fisiológica, mental, económica, cultural ou social, sem haver, no

entanto, referência a um identificador específico como o nome ou um número de identificação.

2.2.2 Dados Sensíveis

O RGPD demonstra grande preocupação com dados pessoais de cariz mais sensível e especial, que permitam uma caracterização detalhada dos titulares de dados, aumentando a abrangência de alguns tipos de dados ou incluindo novos. Estão considerados no Regulamento como sensíveis os dados pessoais como a origem racial ou étnica, as convicções religiosas ou filosóficas, as opiniões e filiações políticas e a filiação sindical, bem como o tratamento de dados genéticos, biométricos, relativos à saúde ou à vida e orientação sexual de uma pessoa.

2.2.3 Entidades intervenientes

O RGPD define o papel das entidades intervenientes. Estes quatro principais intervenientes são: o titular dos dados, como o indivíduo a quem pertencem os dados pessoais; o controlador de dados a quem o titular confia os seus dados pessoais, que determina o propósito e os meios do processamento de dados; o processador de dados, que processa os dados pelo e para o controlador; as autoridades supervisoras, que fiscalizam a proteção de dados em jurisdições específicas.

2.2.4 Licitude de Processamento

Com a entrada em vigor do RGPD, a recolha e tratamento de dados pessoais passa a seguir regras mais rigorosas, podendo acontecer apenas por motivos específicos, explícitos e legítimos, havendo sempre lugar a consentimento por parte do Cidadão. Esta recolha e tratamento de dados pessoais pode ocorrer quando são necessários para a execução de um contrato em que o titular dos dados é parte, ou para diligências pré-contratuais a pedido do titular dos dados. Estão também excluídas da aplicação do RGPD, a recolha e o tratamento de dados pessoais quando para cumprimento de obrigações jurídicas do processador ou controlador dos dados ou se

necessário para o cumprimento de funções de interesse público. Também estão salvaguardados os processamentos de dados pessoais para a proteção dos interesses vitais do titular ou em situações de interesse legítimo.

2.2.5 Consentimento

O Regulamento prevê o fortalecimento do verdadeiro significado da expressão "consentimento de condições" pelo que as organizações deixam de poder utilizar termos e condições longos, cheios de termos técnicos e jargão jurídico, uma vez que o pedido de consentimento deve incluir o objetivo da respetiva recolha e tratamento de dados pessoais e ser escrito numa linguagem clara, inteligível e estar facilmente acessível. A manifestação do consentimento deve ser dada de forma clara, explícita e para um objetivo de processamento concreto, facilmente distinguível de outros eventuais objetivos, que, caso existam, deverão também ser alvo de pedidos de consentimento próprios. Com o RGPD deverá ser possível anular o consentimento, i. e., deverão ser criados mecanismos que possibilitem, por vias legais e claras, a anulação do consentimento. Mais concretamente, deve ser tão fácil anular o consentimento como consentir.

2.2.6 Alcance Territorial Aumentado (Aplicabilidade Extraterritorial)

Uma mudança significativa no cenário regulatório da privacidade de dados vem com a jurisdição alargada do RGPD, uma vez que se aplica a todas as organizações que processam os dados pessoais dos cidadãos que residem na UE, independentemente da localização da organização. Anteriormente, a aplicabilidade territorial da diretiva era ambígua e referia-se ao processamento de dados "no contexto de um estabelecimento". Este tópico surgiu em vários casos judiciais de alto perfil. O RGPD torna a sua aplicabilidade muito clara - o processamento de dados pessoais da UE, independentemente do local do processamento, está abrangido pelo Regulamento. Em casos em que a atividade do cidadão se relacione com a UE (comércio de bens e serviços

e acompanhamento do comportamento que ocorre na UE), os seus dados pessoais estão também abrangidos pela aplicabilidade do RGPD.

2.2.7 Direitos dos Titulares de Dados Pessoais

O Regulamento dá ao titular dos dados um conjunto de direitos sobre o acesso e controlo dos seus dados pessoais e sobre o propósito de tratamento dos mesmos. Indubitavelmente, as mudanças mais relevantes são o Direito ao Acesso e Retificação, o Direito ao Esquecimento, a Portabilidade de Dados e o Processo Automático de Decisão, como abordados nos pontos seguintes.

2.2.7.1 Transparência - Direito ao Acesso e Retificação

Os direitos dos cidadãos contemplados no RGPD, permitem que este obtenha informação sobre os seus dados pessoais, nomeadamente se estes estão a ser processados, onde e com que objetivos. Além disso, o controlador deve fornecer uma cópia dos dados pessoais em formato eletrónico, quando e se solicitado, sendo que o primeiro pedido é obrigatoriamente gratuito. Sempre que solicitado pelo titular, os dados devem ser retificados ou atualizados. Esta é uma mudança significativa para a transparência dos dados e o aumento de direitos dos cidadãos.

2.2.7.2 Direito ao Esquecimento

Também conhecido como *Right to Erasure*, o direito ao esquecimento confere à pessoa em causa, a possibilidade de, junto do controlador de dados, apagar os seus dados pessoais, interromper a disseminação dos dados e, potencialmente, suspender o processamento dos dados fornecidos a *third parties*. As condições da eliminação, conforme descrito no artigo 17º, incluem os dados que não são mais relevantes para os fins originais do processamento ou as pessoas em causa pretenderem retirar o seu consentimento. Deve também notar-se que este direito exige que os controladores

avaliem os direitos dos sujeitos face ao "interesse público na disponibilidade dos dados" ao considerar esses pedidos, uma vez que para determinados serviços é estritamente necessário o armazenamento de informação relativa ao utilizador, como por exemplo, o prestador de serviços necessita de determinados dados pessoais para efeitos de faturação.

2.2.7.3 Portabilidade de Dados

O RGPD introduz nas normas sobre a proteção de dados a noção e as regras para a portabilidade de dados - o direito de um cidadão solicitar, enquanto titular de dados, que os seus dados pessoais sejam transmitidos a si ou a outro controlador. Na prática o controlador "de origem" deverá enviar os dados para o titular ou para um novo controlador num "formato estruturado, de uso corrente e de leitura automática" [14] e eliminá-los, desde que esses dados não sejam necessários para o cumprimento de obrigações (ou outras razões legítimas) por parte do controlador.

2.2.7.4 Processo Automático de Decisão

Este direito dá aos titulares dos dados o direito a não estarem sujeitos a decisões baseadas exclusivamente em processos automatizados que produzam efeitos jurídicos (ou equivalentes) sobre si.

2.2.8 Responsabilidades dos Controladores e Processadores de Dados

O RGPD vem trazer obrigação aos controladores e processadores de dados e seus representantes nomeadamente ao nível da análise e relato das atividades de processamento de dados pessoais, estipulando também sanções para as organizações que não cumprirem o regulamento.

2.2.8.1 *Data Protection Officers*

Com a entrada em vigor do RGPD, as empresas e organizações que processam grandes volumes de dados pessoais ou dados críticos estão obrigadas a nomear um *Data Protection Office* (DPO), que:

- Deve ser nomeado com base nas qualidades profissionais e, em particular, pelo conhecimento especializado em direito e nas práticas de proteção de dados;
- Pode ser um membro da equipa ou um provedor de serviços externos;
- Os detalhes de contacto devem ser fornecidos ao *Data Protection Analyst* (DPA) relevante;
- Devem ser fornecidos os recursos adequados para realizar as suas tarefas e manter os seus conhecimentos especializados e atualizados;
- Deve relatar diretamente ao mais alto nível da gestão;
- Não deve realizar quaisquer outras tarefas que poderiam resultar em um conflito de interesses.

As empresas que não pertençam à UE mas que processam os dados dos cidadãos da UE também terão de nomear um DPO na UE.

2.2.8.2 Registo das Atividades de Processamento

Os processadores e controladores de dados devem conservar registos das suas atividades, incluindo os nomes e contactos dos responsáveis e sempre que for caso disso, corresponsáveis ou representantes pelos respetivos tratamentos. Ou seja, não é suficiente informar o processamento dos dados pessoais de forma adequada, devem manter evidências sobre todos os processamentos efetuados.

2.2.8.3 Avaliação de Impacto

A Avaliação de Impacto sobre a Proteção de Dados (AIPD) é introduzido na legislação enquanto instrumento de responsabilização dos processadores e

controladores de dados. Enquanto processo, uma AIPD descreve “o tratamento, avalia a necessidade e proporcionalidade desse tratamento e ajuda a gerir os riscos para os direitos e liberdades das pessoas singulares decorrentes do tratamento dos dados pessoais” [19].

2.2.8.4 Privacidade por *Design*

A privacidade por *design* enquanto conceito existe há anos, mas apenas agora, com o RGPD, faz parte de um requisito legal. No seu cerne, a privacidade por *design* exige que a proteção de dados pessoais seja uma preocupação desde o início da conceção de sistemas, e não apenas uma preocupação adicional no final do desenvolvimento. Mais especificamente: "O controlador deve (...) implementar medidas técnicas e organizacionais adequadas (...) de forma efetiva (...) para atender aos requisitos deste Regulamento e proteger os direitos das pessoas em questão".

O artigo 23º exige que os controladores mantenham e processem apenas os dados absolutamente necessários para a conclusão de suas funções (minimização de dados). Este artigo exige também que o acesso aos dados esteja limitado apenas aos intervenientes necessários no processamento.

2.2.9 Sanções e Penalizações

As organizações que violem o RGPD podem ser multadas em valores que atingem os 4% do volume de negócios global anual; caso esta percentagem corresponda a um valor inferior a 20 milhões de euros, sendo este o valor máximo de coima a aplicar pela infração. A multa máxima pode ser imposta pelas infrações mais graves, por exemplo, não ter consentimento do cliente para processar determinados dados em determinadas situações ou violar o núcleo dos conceitos de privacidade por *design*. Existe uma abordagem em camadas para multas, por exemplo, uma empresa pode ser multada em 2% por não ter os seus registos em conformidade (artigo 28º), não notificando a autoridade de supervisão e a pessoa em causa sobre uma infração ou não realizando avaliação de impacto. É importante notar que estas regras se aplicam

tanto a controladores quanto a processadores - o que significa que, por exemplo, os serviços na *cloud* não estarão isentos da execução do RGPD.

2.2.10 Notificação de Violação

Sob a observância do RGPD, a notificação de violação tornar-se-á obrigatória em todos os Estados-Membros onde uma violação de dados é suscetível de "resultar num risco para os direitos e liberdades dos indivíduos". Este procedimento deve ocorrer no máximo 72 horas após se ter tomado conhecimento da violação. Esta notificação deve ser feita também aos seus clientes e aos controladores, "sem atraso indevido", depois de terem tomado conhecimento de uma violação de dados.

2.2.11 Transferências Internacionais de Dados

As mudanças judiciais eram, por um lado, já necessárias devido a invalidações de acordos por parte do Tribunal de Justiça da UE; por outro, o RGPD estabelece mudanças, como nas regras "cláusulas contratuais-tipo", nas jurisdições "permitidas", que possuem regras idênticas à UE e em medidas como o "Escudo de Proteção de Privacidade", uma proposta a um novo acordo com os Estados Unidos (EUA). Na prática, as organizações europeias que pretendam transferir dados para fora da UE devem observar se um conjunto de regras são cumpridas no país de destino ("país terceiro") e caso não sejam, as transferências "terão de ser efetuadas mediante a apresentação de garantias adequadas".[3]

[3] https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/what-rules-apply-if-my-organisation-transfers-data-outside-eu_pt (acedido em 16/02/2018)

2.3 Comunicações Seguras

O protocolo de transferência de hipertexto (HTTP) é um protocolo de comunicação utilizado para sistemas de informação distribuídos e colaborativos. Este é a base para a comunicação de dados em toda a *World Wide Web*, sendo o canal para os pedidos das páginas que o utilizador pretende ver e para a resposta enviada pelo servidor com o conteúdo dessas mesmas páginas.

O protocolo seguro de transferência de hipertexto[4] (HTTPS) garante, para além das funcionalidades do HTTP, a integridade e a confidencialidade dos dados transferidos. Na prática, no protocolo HTTPS a troca de informação entre o servidor e o cliente (ou entre servidores) passa por um sistema cifrado – por meio de certificados digitais – garantindo que a informação apenas está disponível para os intervenientes, já que apesar de ser possível a sua interceção, esta não pode ser compreendida.

Para a existência de um sistema cifrado de comunicação é necessária a utilização de protocolos adicionais, como por exemplo o *Secure Sockets Layer*[5] (SSL) e o *Transport Layer Security*[6] (TLS), que fazendo uso de certificados digitais garantem confidencialidade, integridade e autenticidade ao sistema de comunicação, garantindo a proteção dos dados pessoais em transferência.

2.4 *Web Tracking*

Consideram-se como *Web Tracking* todas as práticas através das quais os *websites* recolhem e armazenam informação relativa aos seus utilizadores, criando e interligando padrões de navegação, i.e., sempre que para efeitos de caracterização do perfil dos titulares de dados, os processadores de dados agrupam, selecionam e estudam os dados de navegação. Na ótica do visitante de um *website*, as práticas de *web*

[4] https://tools.ietf.org/html/rfc2660
[5] https://datatracker.ietf.org/wg/ssl/charter/
[6] https://datatracker.ietf.org/wg/tls/charter/

tracking revelam-se nitidamente com fornecimento de conteúdos que se relacionam com as preferências do utilizador [20].

A informação que é recolhida durante a navegação é de interesse para diferentes organizações, como por exemplo:

- Agências publicitárias, como por exemplo, a Google adWords[7], a Bing Ads[8], a 7search[9], ativamente recolhem informações sobre o utilizador, de forma a criar perfis de navegação. Estes perfis são posteriormente utilizados para direcionar anúncios publicitários. Fazendo uso da informação descrita em cada perfil, por exemplo, idade, sexo, páginas visitadas, palavras procuradas, é possível escolher anúncios cuja probabilidade de despertar interesse seja maior. Desta forma, as empresas conseguem aumentar o sucesso das suas campanhas publicitárias.
- As autoridades podem fazer uso da informação recolhida para identificar possíveis infratores e resolver crimes. O uso destas tecnologias é comum para resolver casos de roubo de identidade e fraude de cartão de crédito.
- Empresas que realizam testes de usabilidade em aplicações *web,* recorrem ao histórico de ações para entender a facilidade com que o utilizador resolveu determinada tarefa. Avaliando, por exemplo, o número de *clicks*, o tempo gasto, a trajetória do cursor do rato, torna-se possível inferir a necessidade e o modo de alteração da aplicação *web*.
- Empresas que realizam análises da *web*, como por exemplo, a *Google Analytics*[10] ou a *Piwik*[11], não se focam tanto na criação de perfis de utilização, mas sim, no estudo do desempenho global do *website*. Tendo em conta, por exemplo, o número de visitas, o número de páginas visitadas, o tempo em cada página, a localização geográfica do tráfego, permitindo efetuar melhorias ao *website*, que antes passariam despercebidos.

A obtenção de informação para testes de usabilidade, em caso de consentimento prévio, ou para análise estatística de dados pseudonomizados, não impõe graves

[7] https://adwords.google.com (acedido em 22/2/2018)
[8] https://bingads.microsoft.com (acedido em 22/2/2018)
[9] http://7search.com (acedido em 22/2/2018)
[10] https://google.com/analytics (acedido em 22/2/2018)
[11] https://campaign.piwik.pro (acedido em 22/2/2018)

violações de privacidade. Por outro lado, a obtenção de perfis de navegação com intuito de diferenciar utilizadores, torna-se um tópico crítico no que diz respeito ao conceito de privacidade[16] [21].

2.4.1 Endereço IP

As técnicas de *web tracking* permitem a recolha de diferentes tipos de informação. Nos pontos seguintes são abordados diferentes mecanismos que permitem a identificação do utilizador numa navegação *web*.

O *Internet Protocol*[12] *(IP)* é um protocolo de comunicação usado em redes *Transmission Control Protocol* (TCP)/IP para o encaminhamento de dados entre uma origem e o destino. Para ser possível enviar dados do utilizador para o servidor é necessário que todos os terminais com acesso à *internet* tenham, obrigatoriamente, um endereço IP associado, que funcionará como identificador.

O endereço IP possibilita a criação de um identificador de tráfego, possibilitando a identificação das máquinas numa comunicação. No entanto, este identificador não é sempre fiável para a identificação dos utilizadores, seja pela utilização de redes *Network Address Translation* (NAT) em que vários utilizadores, dotados de endereços IP privados, partilham o mesmo endereço IP público, seja pelo uso de *Dynamic Host Configuration Protocol* (DHCP), que faz com que o endereço IP de uma máquina varie ao longo do tempo, seja pela mobilidade do utilizador, que se liga à internet em diferentes locais com endereços IP diferentes. Não obstante, segundo o RGPD, o endereço IP é considerado um dado pessoal devendo ser tratado com o mesmo rigor que outros dados pessoais, como nome, morada ou outros.

[12] https://tools.ietf.org/html/rfc791

2.4.2 HTTP *Cookies*

O protocolo HTTP é considerado *stateless*, ou seja, não é guardado um estado entre pedidos ao servidor. Sendo assim, por cada pedido, é aberta uma nova conexão entre o utilizador e o servidor.

Com o aumento das funcionalidades dos *websites*, tornou-se necessário criar mecanismos que permitissem controlar a sessão, identificando o utilizador responsável por pedidos subsequentes. Por exemplo: numa plataforma de compras *online*, é necessário registar os artigos que o utilizador adiciona ao carrinho, antes de finalizar a compra; em plataformas, que requerem *login*, é necessário verificar se o utilizador está autenticado. Os controlos de sessão possibilitam a disposição de páginas personalizadas, tendo semelhanças com mecanismos de *tracking*.

Para contornar a falta de estado do protocolo HTTP e implementar o conceito de sessão são utilizadas *cookies*. Uma *cookie* é um pequeno ficheiro de texto, que o servidor armazena no *browser* do servidor e este envia sempre que faz um pedido a esse mesmo servidor.

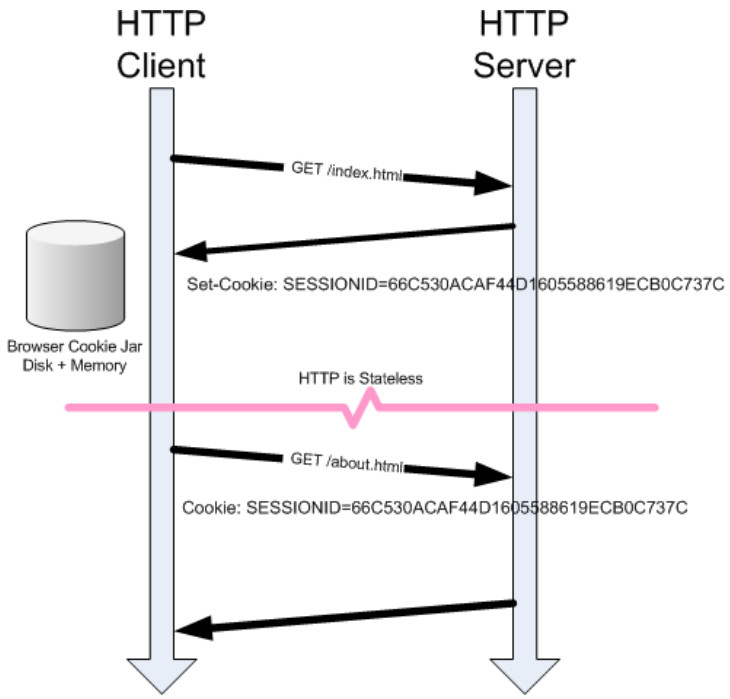


Figura 3 – Representação da interação entre browser e servidor responsável pela imposição de cookies [22]

Quando é iniciado um pedido ao servidor (HTTP *request*), o servidor na resposta a esse pedido pode requisitar a imposição de uma *cookie* (Set-Cookie), que caso seja aceite, é armazenada no browser, que será enviada em cada pedido subsequente a esse mesmo servidor, como ilustrado na Figura 3. As *cookies* são habitualmente utilizadas pelos servidores para controlo de sessão, registo de preferência, para autenticação ou para identificação do cliente; podendo ser distinguidas em dois grupos:

- As *cookies* permanentes, que ficam armazenadas ao nível do *browser*, no equipamento utilizado para o acesso e que são empregues sempre que se volta a visitar o *website* respetivo;
- E as *cookies* de sessão, que são temporárias, permanecendo no *browser* apenas até sair do *website*. A informação recolhida com a utilização destas *cookies* tem como objetivo analisar padrões de tráfego, permitindo identificar *bugs* e outras informações que permitam aos servidores fornecer uma melhor experiência de navegação.

Cada *cookie* tem um domínio associado, correspondente ao domínio do servidor que a colocou. Cada *cookie* apenas pode ser acedida, para leitura ou escrita, pelo respetivo domínio. Ou seja, uma *cookie* do domínio cm-aveiro.pt, não pode ser acedida pelo domínio google.com, e vice-versa. Quando o domínio de uma *cookie* é igual ao domínio visitado, considera-se uma *first party cookie*; quando a *cookie* pertence a um domínio diferente do visitado, denomina-se *third party cookie.*

Para possibilitar o controlo do utilizador entre vários domínios, é necessário que no acesso a múltiplos domínios, sejam também impostas *third party cookies*. A imposição deste tipo de *cookies* é possível por imposição de *scripts*, *iframes*, *forms*, com conteúdo pertencente a um domínio diferente do visitado.

A grande maioria dos processadores de dados utiliza hoje nos seus *websites third parties cookies* quer para monitorizar visitas quer para direcionar publicidade do interesse do seu utilizador. O *Google Analytics* é das ferramentas mais utilizadas na gestão e acompanhamento de visitantes nos sítios na *web* para efeitos estatísticos e de

análise, em parte pela facilidade de integração desta ferramenta, bastando adicionar ao *header* do *website* um código idêntico ao demonstrado na Figura 4.

```
▼<script>
    (function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
    (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
    m=s.getElementsByTagName(o)[0];a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
    })(window,document,'script','//www.google-analytics.com/analytics.js','ga');

    ga('create', 'UA-62545951-2', 'auto');
    ga('send', 'pageview');

  </script>
```

Figura 4 – Screenshot do script do Google Analytics, retirado do código fonte do website do Município de Aveiro (screenshot efetuado a 14/03/2018)

A utilização destes mecanismos no processamento de dados tanto para fins de análise como para fins publicitários permite "conhecer" os utilizadores do *website* e segmenta-los em grupos distintos. Podemos dizer que esta segmentação permite a adaptação dos *websites* e o direcionamento de conteúdos consoante os comportamentos e interesses dos utilizadores; porém esta segmentação pode ser usada de forma discriminatória com base em etnias, filiações partidárias, orientações sexuais e classe económico-social.

2.4.3 *Locally Shared Objects*

O *Adobe Flash* é um *plugin* utilizado para dispor conteúdo interativo nas páginas *web*, por exemplo, vídeos, gráficos, jogos ou animações. O *plugin* tem a capacidade de armazenar *Locally Shared Objects* (LSOs), também designados por *Flash Cookies*. Os LSOs têm funcionalidades semelhantes às HTTP *Cookies* no que diz respeito a controlo de sessão e fornecimento de conteúdos personalizados, podendo ser utilizados, por exemplo, para identificar: o progresso, a pontuação máxima e as preferências de utilização numa aplicação *Flash*.

Em contraste com as HTTP *Cookies*, cuja gestão era da responsabilidade do browser, os LSOs são geridos pelo próprio *Adobe Flash Plugin.* Os LSOs são armazenados no sistema de ficheiros da máquina do utilizador tornando possível identificar o utilizador, mesmo que seja alterado o *browser* de navegação, dado que as

HTTP cookies são armazenadas apenas no contexto do browser. A data de expiração da *Flash Cookie* não é um padrão obrigatório, podendo levar ao armazenamento indefinido de informação desnecessária [23].

No que diz respeito às HTTP *cookies*, os *browsers* disponibilizam opções para a eliminação desse tipo de *cookies* e por outro lado, os LSOs são geridos no *Adobe Flash Player* que tem uma interface própria e independente do *browser*.

2.4.4 *Web Storage*

O *Web Storage* é uma especificação do *Word Wide Web Consortium*, que permite o armazenamento no browser e acesso de grandes quantidades de dados através de *scripts* utilizados no contexto do *browser*. O *Web Storage* é suportado pela maioria dos *browsers* atuais.

As especificações têm dois componentes principais: o *localStorage* e o *sessionStorage*. O primeiro permite o armazenamento persistente de dados, enquanto o segundo é limpo no momento de encerramento do *browser*. À semelhança das HTTP *Cookies*, o *web storage* também funciona com pares de (nome, valor). O envio de dados dos componentes para o servidor é feito com a ajuda de código *Javascript* e é feito na medida do necessário, ou seja, não é enviado em cada pedido, como acontece com as HTTP *Cookies*.

2.4.5 *URL Query Strings*

As *URL Query String* são pedaços de informação que se encontram, habitualmente, no fim do endereço Uniform Resource Locator (URL) e que são enviados para o servidor, possibilitando a criação de identificadores. Por exemplo, o URL http://aaa.com/page.html?userid=123, é enviado para o servidor, como um pedido HTTP GET que posteriormente será analisado possibilitando identificar o utilizador através do *userid*. As restantes páginas acedidas seguem o mesmo procedimento.

Quando mal utilizado, este método possibilita a obtenção de informação sensível, dada a possibilidade de o URL conter informações do utilizador ou do servidor. A análise do tráfego de rede permite então, a obtenção de informação sensível, podendo levar a violações de privacidade. Este tipo de mecanismo é usado, frequentemente, em casos de impossibilidade de utilização de HTTP *Cookies*.

2.4.6 *Browser Fingerprinting*

Para ser possível adaptar a disposição da página às características do dispositivo de navegação, é necessário partilha de informação entre o *browser* e o servidor. Em alguns casos, por exemplo, quando acedemos a uma página *web* utilizando um dispositivo móvel, temos como resposta uma página diferente daquela que teríamos se tivéssemos acedido de um computador portátil. A capacidade de adaptação ao dispositivo, depende da leitura e envio ao servidor de determinadas características do dispositivo, como: resolução de ecrã, sistema operativo, fuso horário, *plugins* instalados [24].

As informações recolhidas, individualmente, não permitem identificar o utilizador. Porém, um cruzamento das muitas e várias informações obtidas permite criar um perfil de navegação – estas informações podem ser o sistema operativo, o browser utilizado, os *add-ons* utilizados no *browser*, entre outras.

2.4.7 Técnicas de Tracking

É possível analisar o comportamento online de um utilizador e são inúmeras as técnicas que o permitem. Nos próximos pontos são explanadas técnicas ou recursos que permitem conhecer os comportamentos online e a recolha de dados pessoais.

2.4.7.1 *Deep Packet Inspection*

O tráfego, entre o utilizador e o servidor, é encaminhado pela infraestrutura do *Internet Service Provider* (ISP). Portanto, o ISP tem acesso ao conteúdo de todos os pacotes transmitidos pelos seus clientes.

Para o bom funcionamento do encaminhamento é necessário analisar, no protocolo IP, a origem e destino do pacote. Porém, tendo em conta que a totalidade do pacote é transmitido na rede do ISP, é possível que seja analisado não só o cabeçalho do protocolo, mas também os seus dados (*payload*). Caso não exista cifragem do pacote, isto é não se está a utilizar o protocolo HTTPS, o ISP tem a possibilidade de percecionar a atividade *online* dos seus clientes, conseguindo obter, por exemplo, páginas visitadas, credenciais, pesquisas e preferências. Por ventura, caso existam mecanismos de cifragem, apesar do conteúdo do pacote ser transmitido na rede, não é percetível.

2.4.7.2 HTTP *Referrer*

O campo *Referrer* é um campo presente no *header* do protocolo HTTP. A sua principal função é identificar o URL de qual foi originário o pedido. Por exemplo, submetendo um pedido, cuja origem é um *link* contido na página www.exemplo.com, para a página www.exemplo1.com, a segunda tem conhecimento que o utilizador proveio da primeira.

Este campo é usado amplamente em estudos de análise da *Web*, já que permite a um observador de tráfego estabelecer sequências, padrões e sessões de navegação do utilizador, sendo importante para, por exemplo, perceber o sucesso de campanhas publicitárias.

2.4.7.3 *Tracking Cookies*

As HTTP *Cookies* podem ser usadas para rastreio da navegação e criação de perfis de navegação e utilização. Para a criação destes perfis, os utilizadores devem ser

identificados e combinados ao longo dos vários domínios visitados. As *First Party Cookies* apenas podem ser acedidas pelo domínio a que o utilizador está a visitar, podendo controlar as páginas visitadas no mesmo domínio, mas não os vários domínios visitados. Como tal a utilização de *first party cookies* não representa um mecanismo de *tracking* de navegação, uma vez que estes apenas fazem *tracking* da navegação do utilizador no seu próprio domínio, seja para controlos de sessão, seja para fornecimento de conteúdos personalizáveis.

O tracking entre múltiplos domínios é possível através da utilização de conteúdo presente na página HTML não pertencente ao domínio visitado, que sempre que é requisitada, é enviado para os servidores desse domínio o endereço da página visitada. Isto permite ficar a saber o histórico de domínios visitados por cada utilizador. A utilização destes conteúdos pode implicar a utilização de *cookies* de terceiros.

Um exemplo deste tipo de mecanismos é o botão *Like* do *Facebook* [26]. O botão *Like* encontra-se presente em inúmeras páginas de diferentes domínios, o que permite ao *Facebook* fazer *tracking* do utilizador entre os vários domínios visitados.

2.4.7.4 *Zombie Cookies*

Com o aumento da preocupação pela privacidade, os utilizadores tendem com maior frequência a remover as *cookies* nos seus browsers. As *Zombie Cookies*, também chamadas de *Super-Cookies*, foram criadas para resistir às remoções.

As *Zombie Cookies* armazenam redundantemente a informação em diversos locais, como: HTTP *Cookies*, *Flash Cookies*, *Silverlight Isolated Storage* e *Web Storage*. Sempre que a informação em uma destas localizações é removida, ela é recriada com o auxílio da informação armazenada nas restantes localizações, através de código *Javascript*.

A *Evercookie*[13] é um projeto *open*-source de Zombie Cookies, desenvolvido em *Javascript* que replica as *cookies* pelos diferentes locais de armazenamento no *browser*.

[13] https://samy.pl/evercookie/ (acedido em 26/2/2018)

2.5 Medição de privacidade na *web*

Nesta secção serão abordadas diversas ferramentas técnicas que possibilitam a automatização da pesquisa de informação que permitem aferir a privacidade e segurança nos sítios na *web*.

Uma infraestrutura que possibilite a automatização da medição de privacidade na *web,* em inglês, *Web Privacy Measurement* (WPM), deve considerar três fases:

- o *input* – simula o utilizador que visita e interage com os *websites*;
- o *output* – regista o comportamento do *website* perante o utilizador, na personalização de conteúdo, o registo das páginas visitadas e cliques efetuados, etc.;
- e a análise – que representa o estudo e interpretação dos resultados obtidos na fase anterior.

A recolha automática de dados que demonstram a conformidade com as normas de privacidade e segurança pode ser feita através da simulação de navegação de utilizadores reais.

Foram consideradas várias opções na determinação da escolha da plataforma a utilizar, desde o desenvolvimento de raiz de uma plataforma especifica para este trabalho até à utilização de outras plataformas: (i) o OpenWPM[14] mereceu desde cedo um destaque pela sua flexibilidade e variedade de opções; (ii) o *FourthParty*[15], um *plugin* para *Firefox*[16] que apesar de permitir a instrumentação não é automatizado e apresenta menos opções que o OpenWPM; (iii) o *AdFisher*[17], uma ferramenta de automatização das visitas e ações do utilizador nos *websites*, focada na comparação da personalização de conteúdo tendo como base diferentes perfis de navegação; (iv) o FPDetective[18], uma plataforma de deteção de *browser fingerprinting*, que utiliza *browsers* leves e medições *stateless*.

[14] https://github.com/citp/OpenWPM (acedido em 26/2/2018)
[15] https://github.com/fourthparty/fourthparty (acedido em 26/2/2018)
[16] https://www.mozilla.org/firefox/
[17] https://github.com/tadatitam/info-flow-experiments (acedido em 26/2/2018)
[18] https://github.com/fpdetective/fpdetective (acedido em 26/2/2018)

Tendo em conta que o *FPDetective* não possui, ao contrário do OpenWPM, a funcionalidade de registo das *cookies*, e que o *Adfisher* e o *FourthParty* podem ser utilizados com o OpenWPM, optou-se por utilizar esta plataforma, como descrito na secção 3.7.

2.6 Mecanismos de defesa

Se por um lado os mecanismos de recolha de dados são diversos e, muitas vezes, complexos para o utilizador; por outro, este último pode defender-se previamente desses mecanismos, ocultando ou mascarando os seus dados – como exemplificado nos pontos seguintes.

2.6.1 *Hiding IP Addresses*

O endereço IP, como referido anteriormente, pode ser utilizado para identificar a fonte e o destino de tráfego. Os administradores dos servidores, que alojam as páginas da *web*, têm capacidade de analisar os registos de *log* nos servidores, possibilitando identificar as origens de tráfego pelo endereço IP. Os ISP têm até a capacidade, através de técnicas de *Deep Packet Inspection*, de analisar o conteúdo de dados reais no fluxo de comunicação. Alguns ISP admitiram vender, esse tipo de registo, a agências publicitárias na *web*.

Apesar de não ser possível identificar o utilizador com base apenas no seu endereço IP, a junção de outro tipo de informação por parte dos controladores de dados, como por exemplo *browser* utilizado, *plugins* instalados e resolução de ecrã, permite a distinção entre utilizadores com o mesmo endereço IP. De forma a evitar mecanismo de *tracking* com base no endereço IP, existem várias soluções:

2.6.1.1 *Proxy Servers*

Proxy servers, também chamados de *proxies*, atuam como intermediários no encaminhamento de tráfego entre o utilizador e o servidor. Na presença de um *proxy server*, quando é requerido um pedido por uma página *Web*, é inicialmente feito o pedido ao *proxy*, que posteriormente reencaminhará o pedido ao servidor que aloja a página *Web*.

Com a utilização de *proxies*, o endereço IP do utilizador só é conhecido pelo próprio *proxy*, enquanto o servidor que aloja a página *Web*, apenas tem conhecimento do endereço IP do *proxy*. Os *proxy servers*, além dos objetivos de anonimato, permitem também: armazenamento em *cache*, filtragem de conteúdo, prevenção de *malware* e impossibilitam a criação de registos de utilização.

2.6.1.2 *Tor and Privoxy*

O *TOR* (*The Onion Router*)[19] é uma rede de túneis virtuais, inicialmente desenhada para proteger comunicações governamentais. Hoje em dia, a rede é aberta ao público, servindo como uma das maiores infraestruturas de anonimização de tráfego – conceito que se assemelha a seguir um caminho confuso e difícil.

Para reduzir as hipóteses de análise de tráfego, sejam análises simples ou sofisticadas, o *TOR* distribui as transações de seus utilizadores por diversas partes da *Internet*, de forma a não ser possível associar destinos e origens a partir de qualquer ponto da rede [32].

Usando o *TOR*, os pacotes de dados não seguem diretamente da origem para o destino. Eles passam por rotas aleatórias e através de vários intermediários que ocultam os passos, conforme ilustrado na Figura 5, visando impedir que um observador em qualquer ponto da rede seja capaz de identificar de onde vieram os dados e para onde eles vão.

[19] https://www.torproject.org/

O *TOR* funciona para fluxos TCP e pode ser usado por qualquer aplicação que suporte *SOCKS*[20]. Por questões de eficiência, o *TOR* reutiliza o mesmo circuito por aproximadamente dez minutos. Requisições posteriores são precedidas pelo estabelecimento de um novo circuito, para evitar que terceiros associem atividades mais antigas a atividades recentes.

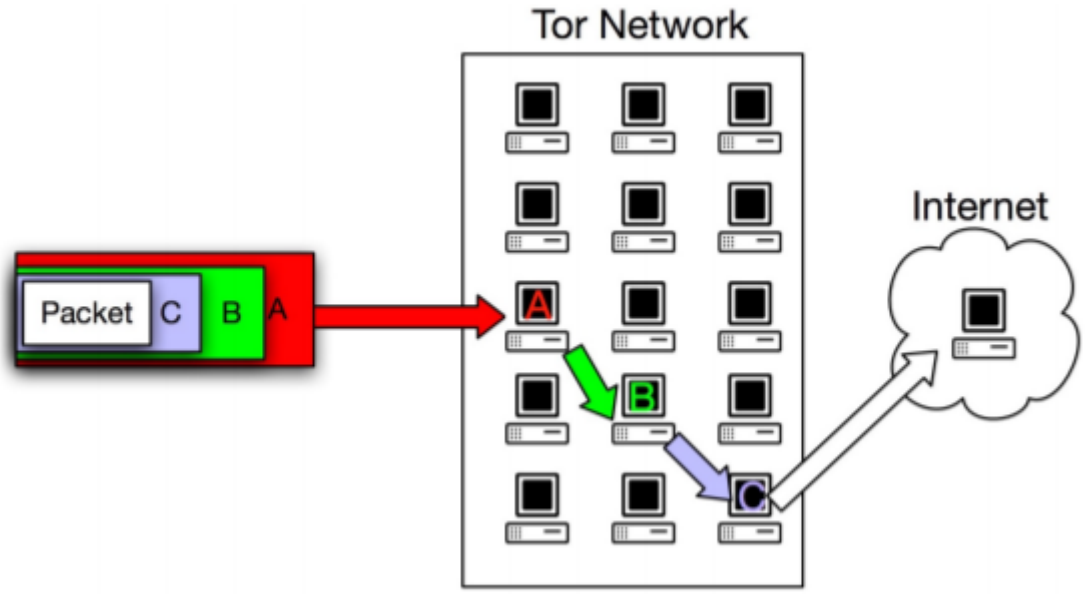


Figura 5 – Um pacote é criptografado uma vez para cada salto na rede TOR. Em cada nó TOR, a camada mais externa de criptografia é removida. [33]

Como semelhança aos *proxy servers* tradicionais, o *TOR* também não possui funcionalidades que permitam impedir mecanismos de *tracking* que não se baseiem em endereços IP, como por exemplo, as HTTP *Cookies*. Por isso, é frequentemente utilizado em conjunto com outros *softwares*, como o *Privoxy*[21].

O *Privoxy* é um *proxy* w*eb* local, sem *cache*, com capacidades avançadas de filtragem para proteger a privacidade, modificando dados da página Web e cabeçalhos HTTP, controlando o acesso e removendo anúncios, impedindo a divulgação de informação não desejada.

[20] https://www.ietf.org/rfc/rfc1928.txt
[21] https://www.privoxy.org/

2.6.2 *Browser Blocking Mechanisms*

A maioria dos *browsers* possibilita mecanismos de combate a *tracking*, seja uma funcionalidade do *browser*, ou uma extensão ao mesmo.

Nos *browsers* atuais, foi implementado um novo modo de navegação. No *Google Chrome*[22] é apelidado de *Incognito mode*, no *Mozilla Firefox* de *Private Browsing* e no *Microsoft Edge*[23] de *InPrivate*. O utilizador, ao fazer uso deste modo de navegação, impede que seja registado o histórico de navegação, sejam acedidas ou impostas *cookies* permanentes, sejam guardadas informações de preenchimento automático e sejam guardadas informações na *cache*.

A utilização deste modo não impede a disposição e acesso de conteúdo pertencente a um servidor diferente do visitado. Algumas das extensões frequentemente utilizadas para lidar com esse problema são *Adblock Plus*[24], *NoScript*[25], *RequestPolicy*[26], *Ghostery*[27] e *HTTPS Everywhere*[28].

2.6.2.1 *Adblock Plus*

O *Adblock Plus* é um *plugin* compatível com o *Mozilla Firefox* e o *Google Chrome.* A sua funcionalidade primária é a remoção de anúncios publicitários indesejados. O *Adblock Plus* funciona com base na criação de listas, contendo expressões regulares, que identificam anúncios no conteúdo HTML. Quando a página é renderizada, é removido todo o conteúdo identificado como anúncio, sendo apresentada a página ao utilizador sem o conteúdo indesejado.

[22] https://www.google.com/chrome/
[23] https://www.microsoft.com/windows/microsoft-edge
[24] https://adblockplus.org/
[25] https://noscript.net/
[26] https://www.requestpolicy.com/
[27] https://www.ghostery.com/
[28] https://www.eff.org/https-everywhere

2.6.2.2 *NoScript*

O *NoScript* é um *plugin* compatível com o *Mozilla Firefox* que seletivamente bloqueia conteúdo executável *Javascript*, *Java Silverlight* e *Flash*. Por definição, todo o conteúdo executável é bloqueado, à exceção do conteúdo permitido pelo utilizador (*whitelist*).

Uma vez que muitos *sites* modernos usam *scripts* por razões legítimas, a abordagem baseada em *whitelist*, implementada pelo *NoScript*, tem desvantagens de usabilidade e requer intervenção frequente do utilizador.

2.6.2.3 *RequestPolicy*

O *RequestPolicy* é um *plugin* compatível com o *Mozilla Firefox* que permite o controlo das conexões a servidores terceiros. Como o *NoScript*, também segue uma abordagem em *whitelist*, sendo que inicialmente todos os pedidos a servidores terceiros são bloqueados. Na Figura 6 está exemplificada a utilização do *RequestPolicy*.

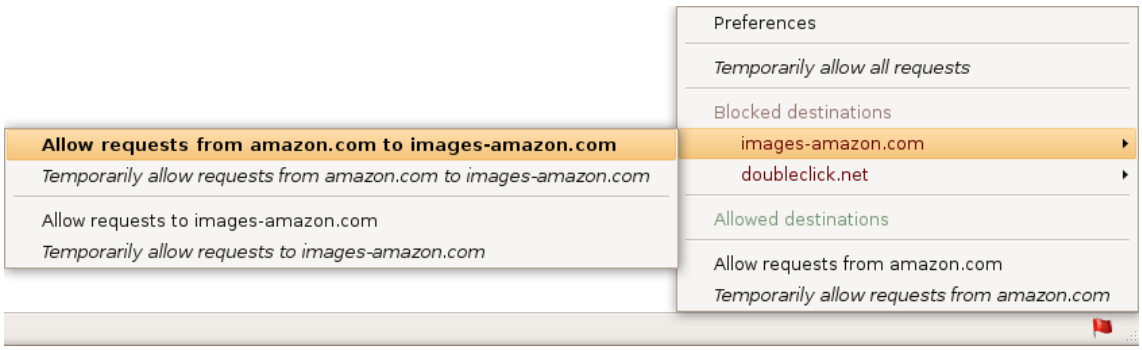


Figura 6 – Screenshot da interface gráfica do plugin RequestPolicy (screenshot efetuado a 20/03/18)

2.6.2.4 Ghostery

O *Ghostery* é um *plugin* compatível com a maioria dos *browsers* atuais: *Edge*, *Opera*, *Firefox* e *Chrome*. É também compatível com dispositivos móveis *Android* e *iOS*. A sua principal funcionalidade é impossibilitar o acesso a conteúdo de servidores terceiros.

O *plugin* é dotado de uma vasta base de dados contendo informação sobre *trackers* de várias fontes, como por exemplo: agências publicitárias, *social media* e agências de análise na *web*. Quando uma página é renderizada, é comparado o conteúdo da página com a informação contida na base de dados, sendo inferido a presença de *trackers*. O *plugin* é personalizável, podendo o utilizador apenas bloquear determinados *trackers*.

Os utilizadores têm de ter cuidado na escolha das fontes para obtenção deste tipo de extensões, já que ocorreram casos de *plugins* réplicas do Ghostery, que, com intuitos maliciosos, colecionavam dados de navegação.

2.6.2.5 HTTPS Everywhere

O HTTPS Everywhere[29] é um *plugin* compatível com *Google Chrome*, *Mozilla Firefox* e *Opera* que tem como objetivo principal a proteção de dados pessoais do utilizador. Esta ferramenta força a utilização do protocolo HTTPS, sempre que este está disponível, i.e., quando determinada visita a um URL retorna um resultado. Quando isto não acontece, o HTTPs Everywhere pode, se assim configurado, permitir a utilização do protocolo não seguro, notificando o utilizador.

Este *plugin* permite também a criação de uma *whitelist* que contempla uma lista de *websites* ou domínios em que não é forçada a alteração do protocolo para HTTPS.

[29] https://addons.mozilla.org/pt-PT/firefox/addon/https-everywhere/ (acedido em 20/3/2018)

2.6.3 *Opt-out cookies*

Algumas empresas que executam mecanismos de *tracking* na *web* oferecem aos utilizadores a possibilidade de recusarem esses mecanismos, configurando uma *cookie* especial (*opt-out*), que é reconhecida e honrada pelos seus servidores negando a existência de futura angariação de dados. Na presença desta *cookie*, é bloqueada a futura imposição de qualquer outro tipo de *cookie*.

As *opt-out cookies* têm desvantagens de usabilidade. Os utilizadores que se preocupam com a sua privacidade e limpem regularmente os seus *cookies*, também podem excluir acidentalmente suas *opt-out cookies*. Além disso, a criação de *opt-out cookies* requer que o utilizador reconheça e encontre o *site* de cada rede de publicidade que deseja excluir.

Existem extensões de navegação, como o *Beef Taco*, que facilitam o gerenciamento de *opt-out cookies*. Outra fraqueza desta abordagem é o desacordo sobre o que um pedido de exclusão implica [34]. Algumas empresas ainda registarão dados de perfil e apenas se absterão de usá-los para anúncios direcionados. As *opt-out cookies* exigem a cooperação do setor para bom funcionamento, i.e., todas as empresas e organizações envolvidas devem honrar as normas estabelecidas.

2.6.4 HTTP DNT *header*

Uma possível abordagem de exclusão alternativa às *opt-out cookies*, é o campo *Do Not Track*[30] *(DNT)* no *header* do protocolo HTTP. Quando configurado pelo *browser*, ele sinaliza, o servidor na *web*, que o utilizador não deseja ser rastreado.

No entanto, semelhante às *opt-out cookies*, os servidores não são de modo algum forçados a honrar o pedido de não fazer *tracking* [35]. Como o cabeçalho só foi introduzido recentemente e porque ainda não existem leis que o regulamentem, a maioria dos servidores da *web* de hoje simplesmente ignoram o campo. Ainda é necessário um quadro regulamentar com mecanismos de execução efetivos. Mas,

[30] https://www.w3.org/TR/tracking-dnt/

mesmo com as leis em vigor, o cabeçalho DNT deve ser usado em combinação com outros mecanismos de defesa, como por exemplo: *opt-out cookies*, *proxy servers* e *browser plugins* [36].

3 Solução Proposta

Neste capítulo será apresentada a solução proposta, o que se pretende que esta registe e analise e quais as tecnologias que permitem averiguar se os *websites* se encontram em conformidade com as normas existentes.

Pretende-se utilizar nesta análise à privacidade do utilizador uma plataforma que integre diversas funções devido à complexidade da análise da conformidade, pela sua amplitude e especificidade.

A plataforma deve ser capaz de analisar a existência de mensagens visuais que informem e/ou recolhem o consentimento para a utilização de *cookies*, a presença de "Políticas de Privacidade", a disponibilização de comunicações seguras, a existência de mecanismos de *tracking* e a eficiência dos mecanismos de defesa.

3.1 Dados a recolher

O objetivo deste trabalho é a análise da conformidade dos sitios na Web das camaras municipais com os requisitos de privacidade previstos na legislação nacional e europeia, nomeadamente os analisados nas secções 2.1 e 2.2. Esta análise de

conformidade é limitada, uma vez que é feita segundo o ponto de vista de um utilizador que acede a um sítio na Web e, por essa razão, não tem em linha de conta nenhuma informação sobre o funcionamento interno das câmaras. Assim, os dados de navegação cuja aquisição foi considerada relevante são os seguintes:

- Aviso sobre utilização de cookies
- A presença de uma página dedicada a divulgar a política de privacidade da câmara municipal
- A implementação de protocolo de comunicação segura (HTTPS) para proteger o acesso ao sítio na Web
- A existência de mecanismos de tracking nos sítios na web das camaras municipais

Estes dados são detalhados nas secções seguintes.

3.2 Informação de utilização de *cookies*

Para permitir uma maior aproximação à conformidade com as normas existentes não é suficiente a existência de mensagens que informem o utilizador da utilização de *cookies*, é necessário oferecer a possibilidade de este aceitar ou recusar explicitamente a referida utilização, tendo também prestado ao utilizador informação relativa aos motivos e à finalidade da utilização de *cookies*. Esta informação surge geralmente sob a forma de um *popup* cujas mensagens podem variar desde a mera informação de utilização de *cookies* até à opção clara de aceitação.

3.2.1 Aviso sobre utilização de cookies

Como foi indicado na secção 2.4.7.3 a utilização de alguns tipos de cookies pode ser usada para finalidades potencialmente prejudiciais para a privacidade dos utilizadores. Assim, de acordo com a legislação europeia das comunicações (Diretiva 2009/136/CE), o utilizador deve ser informado sempre que alguns tipos de cookies são utilizados pelos sitíos, nomeadamente: qualquer cookie de sítios na web de

terceiros (*third parties*) e cookies de longa duração do próprio sítio. Além disso, de acordo com o espírito do RGPD (a lei das comunicações eletrónicas ainda está em discussão e não foi aprovada), além de ser informado, o utilizador deverá poder dar ou negar o seu consentimento para essa utilização. Neste contexto, considerou-se relevante verificar quais os sítios na Web das câmaras municipais, que informam os utilizadores sobre a utilização de cookies, quais as mensagens usadas nesses avisos e que cookies são usados.

Tendo como base os três exemplos das figuras seguintes – modelos gerados com a ferramenta de demonstração da *Cookie Consent by Insites*[31] – podem-se analisar diferentes níveis de informação e aceitação da utilização de *cookies*.

Na maioria dos *websites* a mensagem apresentada ao utilizador é meramente informativa [37], como exemplificado na Figura 7, apresentando geralmente a informação que o *site* utiliza *cookies* podendo ainda identificar a melhoria da experiência de utilização como propósito ou até incluir um *link* para a página de Política de Privacidade. Nestes *popups*, o botão pode conter mensagens como "Ok", "Compreendi" ou até "Fechar", não estando implícito qualquer consentimento por parte do utilizador.

Figura 7 – Popup contendo informação relativa à utilização de cookies, sem requisição de consentimento

Conforme demonstrado na Figura 8, podemos encontrar *websites* em que, com uma mensagem de informação idêntica à anterior, a aceitação da utilização de *cookies* é mais clara, sendo apresentado um botão em que explicitamente o utilizador aceita a utilização destes mecanismos (*opt-in*), embora não seja dada a opção de recusar essa utilização.

Figura 8 – Popup contendo informação relativa à utilização de cookies, com requisição de consentimento

[31] https://cookieconsent.insites.com/ (acedido a 08/04/2018)

Neste último exemplo, a Figura 9 representa uma mensagem que permite ao utilizador aceitar ou recusar (*opt-out*) a utilização de *cookies* nesse *website*. No exemplo apresentado, como na maioria dos *websites*, é dado mais destaque à opção "Aceitar", apresentando a opção que permite recusar a utilização de *cookies* com uma aparência mais discreta, direcionando o utilizador para uma das opções.

Figura 9 – Popup contendo informação relativa à utilização de cookies, com requisição de consentimento ou recusa do mesmo

Da análise anterior depreende-se que cada um destes exemplos representa melhor conduta de atuação que o anterior. Sendo claro que a inexistência de qualquer tipo de mensagem informativa representa, por si só, uma falta de preocupação com a conformidade com as normas e despreocupação com privacidade do utilizador nos seus sítios na *web*, nomeadamente no que toca à garantia do consentimento explícito na utilização de *cookies*.

Pretende-se por isso que a plataforma registe a existência destas mensagens nas páginas das Câmaras Municipais a fim de averiguar qual o nível de consentimento pedido ao utilizador.

Para garantir que os websites estão em conformidade com as normas é necessário, no primeiro acesso ao *website* de um município, recolher uma captura de ecrã, e posteriormente analisar o conteúdo das mensagens apresentadas.

3.3 Presença de "Políticas de Privacidade"

Representando uma forma consensual de informação sobre o registo e o tratamento de dados, a página de "Políticas de Privacidade" é um mero indicador para a conformidade com o RGPD, uma vez que a sua existência não é por si só suficiente para garantir a conformidade, dado que, o conteúdo destas políticas deve, também, ser tido em conta.

Com a utilização destas páginas cada município tem a oportunidade de disponibilizar e centralizar as informações relativas à utilização de dados pessoais dos visitantes dos seus *websites*. Nesta página a entidade responsável pelo *website* deve definir de forma clara, concisa, transparente e inteligível, as práticas através das quais os dados pessoais são recolhidos, quais os dados recolhidos e a finalidade dessa recolha. Devendo ser de fácil acesso, nesta página deve ainda existir a informação relativa à partilha de informação com terceiros e a possibilidade de objeção à recolha e tratamento de dados.

Como vimos atrás, são muitas as informações que estes manifestos podem e devem conter, assim, o processo de automatização apresenta como lacuna a não avaliação do conteúdo destas páginas. Contudo, a existência de Políticas de Privacidade representa, por si só, uma preocupação da entidade em garantir alguma transparência no tratamento de dados.

A interpretação textual do conteúdo destas páginas leva à necessidade da criação de um modelo complexo que processe informação de natureza tão variada e subjetiva, pelo que, a implementação desta funcionalidade implicaria um estudo bastante detalhado podendo requerer um *timing* próprio e um projeto independente da plataforma a utilizar.

3.4 Disponibilização de Comunicações Seguras

A análise à utilização dos protocolos HTTP e HTTPS pelos *websites* dos municípios permite perceber a preocupação que determinada entidade demonstra com a proteção dos dados pessoais dos seus utilizadores. Na navegação *web*, cada entidade pode optar por redirecionar o tráfego forçando a utilização do protocolo HTTPS, noutros casos, não ocorre o direcionamento ou não existe sequer este protocolo seguro. Neste trabalho será verificado se o protocolo HTTPS é priorizado em relação ao protocolo HTTP.

Na maioria dos *websites*, é requerida grande quantidade de informação, por exemplo, através da utilização de formulários; nestes, dependendo da finalidade do respetivo formulário, são pedidas informações de natureza variada, como nome, dados de contacto, dados de identificação civil ou fiscal, entre outros. A transferência destes dados é especialmente sensível em navegações que utilizem o protocolo HTTP, uma vez que os dados não são cifrados e podem estar visíveis a terceiros interessados [38].

A utilização do protocolo HTTPS permite também garantir a autenticidade da página *web* visitada, dada a utilização de certificados digitais emitidos por entidades fidedignas. Com a utilização de certificados fidedignos é dada a garantia ao utilizador que o *website* é efetivamente propriedade do respetivo município.

3.5 Existência de Mecanismos de Tracking

Como verificámos anteriormente existem diversos mecanismos de *tracking* que podem ser utilizados nos *websites*. Pretende-se que na análise sejam registados: o *tracking* ativo (*cookies*, LSO, *Web Storage*, IS, entre outros), o *tracking* passivo (*browser fingerprinting*) e o fluxo de dados entre *trackers*.

Com a cedência de dados a terceiros, através por exemplo das *Third Party Cookies*, podem ser postos em causa diversos aspetos da privacidade do utilizador. A partilha destes dados (como localização geográfica, características da máquina, *browser* utilizado, *plugins* instalados, entre outros) permite criar perfis de navegação, permitindo identificar o terminal da navegação criando uma correlação entre este terminal e os *websites* visitados. Através do histórico de navegação é possível identificar os interesses, gostos e escolhas do utilizador - esta informação como referido no ponto 2.4, é útil para diversas entidades pois permite a segmentação de utilizadores e consequentemente a personalização de conteúdos. Este processo é geralmente efetuado sem um conhecimento e consentimento explícitos do utilizador e de forma pouco transparente.

3.6 Eficiência dos Mecanismos de Defesa

Como consequência da análise à existência de mecanismos de *tracking*, considera-se importante aferir a eficiência dos mecanismos de defesa na navegação nos *websites* dos municípios, uma vez que é através destes que o utilizador pode, de forma pró-ativa, proteger os seus dados.

Neste trabalho irão ser testados dois mecanismos de defesa de natureza distinta: um sendo uma implementação a nível do protocolo de comunicação, o *Do Not Track*; e outra como uma extensão ao *browser*, o Ghostery.

3.7 Plataforma a utilizar

A plataforma concebida de raiz para este trabalho permite a análise dos diversos *websites* dos municípios e foi desenvolvida em *Python*, utilizando como ferramenta de automatização o *Selenium*[32]. A plataforma visita as páginas *web* de cada município registando, numa Base de Dados MongoDB[33], informações da navegação, tais como, *cookies* e suas propriedades – datas de criação e de expiração, se é *cookie* de sessão ou não, entre outros –, a presença de *scripts Javascript* e lista os *links* presentes em cada página, analisando se estes se referem a páginas de "Política de Privacidade".

Uma vez que se iniciou este desenvolvimento em simultâneo com o aprofundamento teórico e a análise ao estado da arte, percebeu-se, logo numa fase precoce, que as respostas que se pretendiam obter com esta ferramenta correspondiam praticamente na totalidade às obtidas com a utilização do OpenWPM, que é analisado na secção seguinte – quer ao nível das funcionalidades quer no que toca à demonstração de resultados.

[32] https://www.seleniumhq.org/
[33] https://www.mongodbcom/ (acedido em 11/04/2018)

Analisou-se também o *FPDetective* que, em comparação com o OpenWPM, se destaca pelas funcionalidades de obtenção de dados relativos a *browser fingerprinting*, que é um dos pontos importantes a observar nesta análise; no entanto o *FPDetective* não permite registar informações sobre *cookies* e não permite guardar o estado entre comunicações HTTP – impedindo a análise da obtenção de informação na navegação entre múltiplos municípios, por parte de um mesmo *tracker*.

Sendo que a ferramenta em desenvolvimento necessitaria de melhorias, correções e testes e o *FPDetective* possuí, em comparação, menos funcionalidades, entendeu-se utilizar como base deste trabalho o OpenWPM, que possui um vasto histórico de trabalhos bem sucedidos [39]. Adicionalmente, desenvolveram-se extensões para o OpenWPM que adicionaram as funcionalidades necessárias para obter os dados pretendidos.

3.7.1 OpenWPM

A plataforma OpenWPM é uma ferramenta flexível e *open-source*, que é desenvolvida em *Python*. O OpenWPM foi concebido com o intuito de permitir a *developers* e investigadores efetuar medições confiáveis com *browsers* reais, garantindo estabilidade, modularidade e escalabilidade. A plataforma fornece *browsers* extensíveis e instrumentalizados como instâncias isoladas, suportando essencialmente qualquer experiência de WPM.

A plataforma para medição automatizada de privacidade na *web* possui três componentes – simulação de utilizadores, experiência de navegação e registo de informação (*cookies*, LSO, entre outros).

3.7.1.1 Funcionalidades

O OpenWPM deve responder a um conjunto específico de necessidades e devem ser definidos previamente alguns critérios e opções. Nos três pontos seguintes estão

descritas estas considerações prévias que possibilitam a correta recolha de resultados, e na forma como estes interagem e reagem aos seus visitantes.

3.7.1.1.1 Simulador de Navegação

No desenvolvimento do OpenWPM foram consideradas uma variedade de opções que permitissem a simulação de navegação, i.e., instruir o *browser* para visitar um determinado conjunto de páginas e executar um determinado conjunto de ações. As opções consideradas dividem-se em três categorias: as bibliotecas HTTP – como *curl* ou *wget* – *browsers* leves – como o PhantomJS[34], uma implementação originária do *WebKit*[35] – e *browsers* completos – como o *Google Chrome* ou o *Mozilla Firefox*.

Na criação do projeto OpenWPM desconsiderou-se a utilização de bibliotecas HTTP – por não ser possível a execução de código *Javascript* – e os *browsers* leves – por não oferecerem suporte a *plugins* e poderem possuir leves diferenças decorrentes da falta de mecanismos para a execução de código *Javascript* mais atual (ECMAScript 2015). Optou-se pela utilização do *Selenium* por possuir um *driver* comum a diferentes *browsers* e permitir a utilização de *plugins*, o que possibilita a obtenção de resultados mais variados e coerentes, simulando de forma mais real a interação do cidadão com o município [28]. Nesta análise à privacidade do utilizador em sítios *web* dos municípios optou-se pela utilização do *browser Mozilla Firefox*.

3.7.1.1.2 Instrumentalização do *Browser*

Na realização do estudo da privacidade do utilizador em sítios na *web* dos municípios portugueses, é necessário recolher informação com origem variada. A recolha tem de abranger o acesso a diferentes componentes do *browser*, sendo necessário registar todas as chamadas a esses mesmos componentes, por exemplo,

[34] http://phantomjs.org/
[35] https://webkit.org/ (acedido em 28/2/2018)

leitura ou escrita de elementos do s*essionStorage*, *localStorage*, *navigator*, *plugins*, entre outros.

A plataforma utilizada permite qualquer opção de instrumentalização - modificar o código-fonte do *browser* diretamente [29] [30], utilização de um *plugin, por exemplo, o Ghostery,* ou até mesmo direcionar tráfego por um *proxy*, por exemplo, com o auxílio do *FoxyProxy*[36]. Esta capacidade de instrumentalização do *browser* permite a referida modularidade do OpenWPM, permitindo alterações ou extensões de forma facilitada.

3.7.1.1.3 Predicados da Plataforma

O OpenWPM permite o aumento da velocidade por meio de paralelismos do *browser* garantindo a capacidade de instâncias diferentes poderem realizar acessos concorrentes à mesma base de dados. A principal vantagem da automatização dos *browsers* é permitir navegação a diferentes *sites* repetidamente a uma taxa inviável para seres humanos.

De forma a garantir o rigor científico, o OpenWPM permite registar dados num formato padronizado possibilitando a fácil execução dos *scripts* de análise para verificar os resultados.

3.7.1.2 *Design* e Implementação

Ainda que baseado em diversas plataformas já existentes, o OpenWPM possui diferenças fundamentais ao nível do *design* oferecendo suporte a medições modulares, abrangentes e de manutenção simples. Por exemplo, ao contrário do FPDetective, o OpenWPM suporta medições *stateful* [29]. Estas medições – com informação de estado – são importantes para o estudo do ecossistema de rastreio: quando a navegação na *web* é *stateless*, cada visita a uma determinada página é sempre vista como um novo

[36] https://addons.mozilla.org/pt-PT/firefox/addon/foxyproxy-standard/ (acedido em 28/2/2018)

utilizador. Em contraste, a plataforma utilizada permite registar estado entre pedidos, possibilitando estudar o reaparecimento de *cookies e* simular perfis de utilizadores reais.

A infraestrutura de recolha de dados e automatização de *browsers* está dividida em três módulos principais: gestores de *browsers* (*browser manager*) que atuam como uma camada de abstração para automatizar instâncias de *browsers* individuais, um gestor de tarefas (*task manager*) voltado para o utilizador que permite distribuir comandos aos gestores de *browsers* e um agregador de dados (*data aggregator*), como uma camada de abstração para a documentação do *browser*. A interação com o gestor de tarefas faz-se por meio de uma linguagem extensível e de alto nível, para controlar a instância do *browser*. Toda a plataforma é construída utilizando bibliotecas *Python*.

Na Figura 10 [31] apresenta-se um esquema sobre a infraestrutura e o *workflow* do OpenWPM. O gestor de tarefas monitoriza os gestores de *browser*, que convertem comandos de alto nível em ações automatizadas do *browser*. O agregador de dados recebe e pré-processa os dados da instrumentalização.

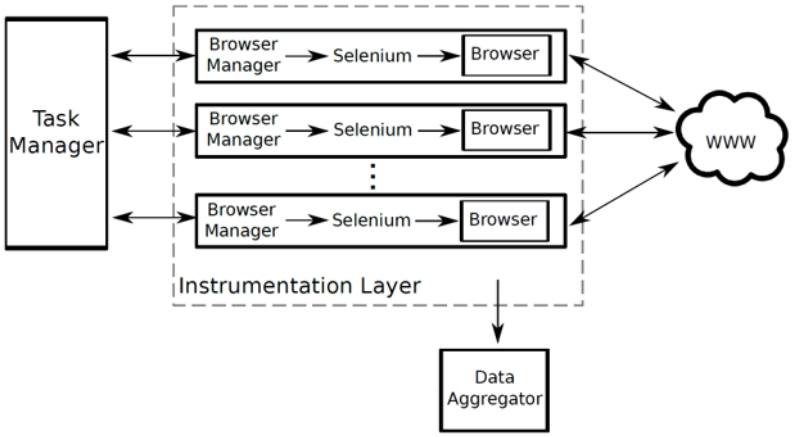


Figura 10 – Arquitetura detalhada do projeto OpenWPM [31]

3.7.1.2.1 Gestores de *browsers*

Uma vez que a ferramenta escolhida para automatização foi o *Selenium*, torna-se necessária a existência de gestores de *browsers* como camada de abstração. Os

gestores de *browser* têm como responsabilidade converter os comandos da plataforma de alto-nível em sub-rotinas do *Selenium*, por exemplo, traduzir o comando visitar determinado *website* numa ação concreta.

O *Selenium* apesar de excelente pela sua operabilidade com diferentes *browsers*, apresenta problemas no que toca à execução de múltiplas instâncias. Neste sentido, cada gestor de *browser* instância uma instância do *Selenium* com um conjunto de configurações específicas de um determinado utilizador, por exemplo, ativar o DNT, o Ghostery ou o AdBlock Plus.

O gestor de *browser* tem a capacidade de iniciar múltiplas instâncias do *Selenium*, porém esta funcionalidade implica grandes consumos de memória e processamento. Para combater esta limitação, o gestor de *browser* possuí também a opção de iniciar várias instâncias do *Selenium* sem representação gráfica. A remoção do componente gráfico permite não só diminuir o consumo de memória e processamento, mas também a diminuição no tempo de recolha de informação.

3.7.1.2.2 Gestor de tarefas

O gestor de tarefas fornece uma interface de linha de comando programável para controlar em simultâneo vários *browsers*. Este módulo do OpenWPM monitoriza os gestores de *browsers*, reiniciando-o se detetar um erro ou não conseguir concluir um comando dentro de um limite de tempo, previamente definido.

A interface de programação da aplicação (API) do gestor de tarefas abre quatro opções para emitir comandos para vários gestores de *browser*: seja por "ordem de chegada", seja para gestores de *browsers* individuais ou para todos os *browsers* em simultâneo – quando o comando é enviado a todos os *browsers*, pode ocorrer de forma síncrona ou assíncrona. Isto permite iniciar facilmente medições em vários *browsers* que escalam para os recursos de sistema disponíveis, especificando apenas o número de instâncias paralelas a serem executadas.

3.7.1.2.3 Agregador de dados

O agregador de dados, que ocorre num único processo isolado, recebe dados durante o processo de análise, manipula-os e grava-os numa base de dados central SQLite[37]. O acesso ao agregador de dados é efetuado através de uma interface que pode ser facilmente ligada a um qualquer número de gestores de *browsers* ou processos de instrumentalização. Além disso, isolar o agregador de dados garante que quando o *browser* para noutros processos, não corrompe nenhum dado nem nega o acesso a outros processos de recolha de dados. Em comparação com as bases de dados *client-server*, as bases de dados SQLite locais pela sua estrutura e por não necessitarem de instalação no utilizador final permitem a partilha dos dados, mesmo com outras estruturas, como MySQL ou bases de dados na *cloud*.

3.7.1.2.4 Camada de Instrumentalização

Na conceção do OpenWPM foram consideradas diferentes opções de instrumentalização, incluindo o FourthParty, vários *proxies* e *plugins* personalizados.

Em última análise, compararam-se o FourthParty e o Mitmproxy[38] percebendo de qual das ferramentas se poderia retirar mais valor e funcionalidades: ambas permitem intercetar o tráfego HTTP e HTTPS, possibilitando registar as comunicações HTTP e o seu conteúdo, *scripts* executados e *Javascript* API *calls*; por outro lado e ao contrário do FourthParty, o Mitmproxy apresenta a capacidade de, através de *scripts* em Python, alterar o conteúdo das comunicações HTTP- desta forma, é possível um maior controlo na interação *web*.

Pese embora todas as opções consideradas tenham aspetos positivos e negativos, o Mitmproxy ficou definido como principal meio de instrumentalização do *browser*.

[37] https://www.sqlite.org/ (acedido em 5/3/2018)
[38] https://github.com/mitmproxy/mitmproxy (acedido em 5/3/2018)

4 Implementação

Neste capítulo serão descritas e fundamentadas as opções tomadas e os desenvolvimentos necessários de forma a garantir a análise da conformidade com as normas de privacidade. Estas implementações, tanto as configurações necessárias ao OpenWPM como as extensões desenvolvidas, são abordadas nos itens seguintes, divididas por temáticas.

4.1 *Websites* dos Municípios

A recolha dos endereços dos *websites* dos municípios portugueses foi efetuada recorrendo à Lista de Municípios disponibilizada pela Associação Nacional de Municípios Portugueses (ANMP), por se considerar que esta entidade apresenta as competências julgadas necessárias para listar os municípios e respetivos *websites*.

Foi desenvolvido um *script* em *Python* que faz uso do *Selenium* para obter o conteúdo, disponibilizado como exemplificado na Tabela 1. Posteriormente, perante a

necessidade de recolher apenas informação útil (colunas "Município" e "*Web*") optou-se por utilizar uma biblioteca *Python* (HTMLParser[39]) com o intuito de identificar essa informação através da análise das *tags* HTML.

Esta informação é armazenada numa estrutura JSON[40] para fácil integração com o OpenWPM. Esta informação está listada no Anexo A.

Município	**Presidente**	**Endereço**	**Tel/Fax**	**E-mail**	***Web***
Abrantes	MARIA DO CÉU ALBUQUERQUE	Praça Raimundo Soares 2200-366 ABRANTES	Tel: 241 330 100 Fax: 241 330 190	geral@cm-abrantes.pt	http://www.cm-abrantes.pt
Águeda	JORGE ALMEIDA	Praça do Município 2 3754-500 ÁGUEDA	Tel: 234 610 070 Fax: 234 610 078	geral@cm-agueda.pt	http://www.cm-agueda.pt
Aguiar da Beira	JOAQUIM ANTÓNIO BONIFÁCIO	Av. da Liberdade 3570-018 AGUIAR DA BEIRA	Tel: 232 689 100 Fax: 232 688 894	geral@cm-aguiardabeira.pt	http://www.cm-aguiardabeira.pt
Alandroal	JOÃO MARIA ARANHA GRILO	Praça da República 7250-116 ALANDROAL	Tel: 268 440 040 Fax: 268 440 042	geral@cm-alandroal.pt	http://www.cm-alandroal.pt
Albergaria-a-Velha	ANTÓNIO AUGUSTO LOUREIRO SANTOS	Praça Comendador Ferreira Tavares 3850-053 ALBERGARIA-A-VELHA	Tel: 234 529 300 Fax: 234 522 225	geral@cm-albergaria.pt	http://www.cm-albergaria.pt
…	…	…	…	…	…
Vouzela	RUI MIGUEL LADEIRA PEREIRA	Alameda D.Duarte de Almeida 3670-250 VOUZELA	Tel: 232 740 740 Fax: 232 771 515	geral@cm-vouzela.pt	http://www.cm-vouzela.pt

Tabela 1 – Excerto de tabela da ANMP com os contactos de cada município[41]

4.1.1 Análise a Subdomínios

No âmbito do estudo da privacidade do utilizador nos *websites* dos municípios – e porque se considerou que não era possível uma verdadeira análise sem avaliarmos todo o conjunto de páginas *web* dos municípios –, era espectável que a plataforma fosse capaz de, de forma direta, listar para cada domínio os respetivos subdomínios, para

[39] https://docs.python.org/2/library/htmlparser.html (acedido em 13/04/2018)
[40] https://www.json.org/ (acedido em 13/04/2018)
[41] http://www.anmp.pt/munp/mun/mun101l1.php?cod=20140110 (acedido em 13/04/2018)

que fosse avaliada também a conformidade de cada uma das páginas desses subdomínios.

Pela especificidade das páginas *web* dos subdomínios dos municípios, estas podem recorrer à utilização de mecanismos distintos de *tracking*. Se pretendermos, por exemplo, avaliar corretamente o cumprimento das normas de privacidade e segurança em sítios na *web* no Município de Aveiro temos não só que avaliar todas as páginas *web* do domínio cm-aveiro.pt, mas também as dos subdomínios como bibria.cm-aveiro.pt, mca.cm-aveiro.pt, catalogo.cm-aveiro.pt, agenda.cm-aveiro.pt entre outros.

Foi desenvolvida uma extensão ao OpenWPM para esse efeito que visita uma página aleatória de cada município, retirando diretamente do código-fonte da página HTML todos as referências a *links* nela contida; cada uma destas referências é comparada com o domínio principal do respetivo município e é avaliado se o domínio está contido no *link* referenciado; caso esteja, é pesquisado o texto que precede o domínio e registado, sendo previamente validado se este é único. Os links considerados válidos são guardados numa estrutura JSON que servirá de base para a lista de páginas a visitar.

4.2 Informação de utilização de *cookies*

Para ser possível garantir que, quando existe, se obtém um *screenshot do popup* que apresenta a mensagem informativa sobre a utilização de *cookies*, considerou-se registar a imagem da primeira página visitada para cada município, uma vez que se pressupõe que numa primeira visita ao site de um município deve ser comunicada ao utilizador a utilização de *cookies*; também se registará a representação gráfica de uma outra página visitada, uma vez que, numa primeira fase de análise manual, se percebeu que este *popup* não surge muitas vezes na primeira página sem que haja alguma ação do utilizador (optou-se por registar a quarta página visitada).

No OpenWPM utilizou-se o *pyvirtualdisplay*[42], um *wrapper* para a interface Xvfb[43], que projeta uma representação gráfica do *browser* exibida num *buffer* como um conjunto de *frames*, que posteriormente são agrupados numa só imagem, permitindo assim uma fácil criação de *screenshots* dos *websites* renderizados pelo *browser*. Estas imagens são guardadas no sistema de ficheiros, sendo analisadas posteriormente conforme indicado na secção 5.

4.3 Presença de "Políticas de Privacidade"

Como referido anteriormente, o utilizador de determinado *website* deve estar criteriosamente informado da existência e finalidade da recolha e tratamento dos seus dados pessoais. Esta informação, que deve ser inequívoca e expressamente conhecida e aceite pelo utilizador antes mesmo da recolha, deve continuar a estar acessível e incluir determinados requisitos, como a via através da qual o utilizador pode, por exemplo, corrigir os seus dados ou recusar a recolha e tratamento dos mesmos.

Sabe-se que, usualmente, esta informação é transmitida sob a forma de uma página exclusiva para esse fim, geralmente, designada por "Política de Privacidade"; estas páginas apresentam, dentro de determinadas áreas de negócio ou atuação, conteúdos genericamente idênticos [37].

Conforme indicado na secção 3.3, esta automatização do processo de análise registará apenas a existência de páginas onde constem as Políticas de Privacidade nos *websites* dos municípios, não sendo analisado o conteúdo das mesmas. Neste sentido desenvolveu-se uma extensão ao OpenWPM que permite, durante o processo de análise a cada município, em cada página *web* visitada: registar a ocorrência dos termos "política", "privacidade", "segurança", "cookie", "privacy" e "policy"; registar o contexto em que estes termos ocorrem, i.e., a frase ou parágrafo. Este registo acontece apenas quando estes termos são encontrados num *link*, ficando registado na base de dados *SQLite*, na tabela *privacy_policy_links*, o endereço desse mesmo *link*. Para tal, o gestor

[42] https://pypi.org/project/PyVirtualDisplay/
[43] https://pypi.python.org/pypi/xvfbwrapper (acedido em 14/04/2018)

de *browser* instrui o *Selenium* a analisar em cada página visitada as *tags* HTML, procurando a *tag* <a> e comparando o texto desse elemento HTML com as palavras-chave previamente identificadas.

4.4 Disponibilização de Comunicações Seguras

Para avaliar a utilização de protocolos de comunicações seguras utilizou-se como *addon* ao *Mozilla Firefox* o HTTPS Everywhere, que força o *browser* a utilizar o protocolo HTTPS nas comunicações *web*. Quando é visitado um *link* com HTTPS e a análise retorna resultados, considera que esse município utiliza o protocolo seguro; se, por outro lado, a utilização de HTTPS não retorna uma página válida ou não retorna *links* do respetivo domínio, um mecanismo de *fallback* inativa a utilização do HTTPS Everywhere e visita as páginas recorrendo ao protocolo HTTP – esta ocorrência é registada na base de dados *SQLite* (nas tabelas *http_requests, http_redirects e http_responses*) e é, então, entendido que esse município não disponibiliza um canal de comunicação segura.

Como complemento, foi implementado um *script* adicional externo que verifica a priorização da utilização de um canal seguro, desenvolvido em *Python*, que deliberadamente solicita um canal de comunicação HTTP e verifica a existência de redirecionamento para um canal de comunicação segura. Esta informação é registada numa estrutura *JSON*.

4.5 Existência de Mecanismos de Tracking

Na análise ao *tracking* ativo e *tracking* passivo, com o OpenWPM, instruiu-se o gestor de *browser* a instanciar uma instância distinta do *Selenium* para cada conjunto de páginas a visitar num domínio. Esta definição permitiu que para os diferentes

municípios o estado inicial do *browser* fosse o mesmo. Ao possibilitar esta distinção conseguiu-se garantir resultados independentes para cada município.

Simulando uma primeira visita de um utilizador real ao *website* de cada município, os gestores de tarefas transmitem aos gestores de *browsers* a informação sobre o *link* inicial a visitar, com base na informação identificado Anexo A, ou seja, é enviado ao *browser* um comando sinalizando o pedido de obtenção de uma determinada página *web*. Este comando permite receber o conteúdo da página pretendida e registar todas as comunicações HTTP – *requests, redirects* e *responses* – originários do pedido inicialmente efetuado; a informação é registada numa base de dados *SQLite,* nas tabelas *http_requests, http_redirects e http_responses*.

Nesta primeira página visitada, que como referido é previamente determinada, são recolhidos os *links* que nela existem, através da pesquisa da *tag* <a>, e escolhe o *link* seguinte a visitar, repetindo o processo até ter analisado 50 páginas, por se ter considerado que para a análise a efetuar, não traria valor acrescentado uma análise mais extensiva. Em cada processo de escolha do *link* seguinte a visitar é validado se este pertence ao domínio do município em análise e se já foi visitado anteriormente; assim, a análise poderá terminar antes de ter percorrido as 50 páginas se e quando todos os *links* existentes tiverem sido visitados.

Tendo o gestor de *browser* controlo absoluto de toda a atividade do *browser* é possível recolher todas as informações provenientes da interação com o *website* do município. Nesta interação pode ocorrer a presença de *First* e *Third Parties Cookies*, pelo que se configurará o OpenWPM para que, nestes casos, registe informações como datas de criação e expiração, data do último acesso, o respetivo domínio, o nome e o valor. Estas *cookies* podem ser originárias do protocolo HTTP – lidas através de uma extensão ao *browser*, o Mitmproxy[44], podem ser *Javascript* – analisando o conteúdo dos *scripts* executados e verificando acessos ao objeto *document.cookie* – ou podem ser do *Flash Player* – através da análise ao conteúdo escrito por este no sistema de ficheiros, registando esses mesmos ficheiros.

A verificação de existência de *browser fingerprinting* é efetuada através da análise aos *scripts* executados quando a página é renderizada, registando acessos a

[44] https://mitmproxy.org/

determinadas APIs com potencial interesse de *fingerprinting*. Este registo é possível através da análise integral do *script*, verificando a ocorrência de objetos do *browser*, como, *navigator.plugins*, *window.sessionsStorage*, *window.screen,* e *userAgent*. Destas ocorrências são registados (na tabela *javacript* da base de dados SQLite) vários tipos de propriedades, como por exemplo, data de ocorrência, tipo de operação (leitura ou escrita) e o valor da operação.

Considerou-se importante para este trabalho, efetuar, após esta primeira análise, uma análise aos mecanismos de *tracking* nas visitas consecutivas aos vários domínios dos municípios utilizando a mesma instância do *browser*, em oposição à análise anterior, que utiliza uma nova instância para cada município. Desta forma, considera-se ser possível observar a obtenção de informação na navegação entre múltiplos municípios, por parte de um mesmo *tracker*, registando o acesso a informações comuns aos diferentes municípios.

4.6 Eficiência dos Mecanismos de Defesa

Após a análise à existência de mecanismos de *tracking*, procede-se à análise da atuação dos mecanismos de defesa escolhidos na navegação nos *websites* dos municípios.

De forma a garantir uma análise viável e com resultados confiáveis, julgou-se necessário, para cada mecanismo de defesa, fazer uma leitura sem quaisquer mecanismos de defesa e uma análise utilizando esse mecanismo, num espaço temporal aproximado, i.e., garantindo que cada página era idêntica em cada par de análises, sem que tivessem ocorrido atualizações aos conteúdos das páginas.

Nas análises sem utilização de mecanismos de defesa recolhem-se todas as informações referidas na secção anterior, para serem utilizadas como referência.

Na análise à eficiência do DNT na navegação nos *websites* dos municípios, configurou-se o OpenWPM para utilizar o *Mitmproxy*, de forma a garantir que, em cada pedido HTTP, o *header* DNT estava presente e ativo. Para analisar a eficiência do

Ghostery, instrui-se o gestor de *browser* para utilizar este *plugin* na navegação, com base nas configurações padrão, que incluem determinados filtros predefinidos.

5 Resultados

Neste capítulo serão demonstrados e analisados os dados recolhidos no âmbito deste estudo. Estes dados permitem uma visão sobre a utilização de mecanismos que comprometem a privacidade do utilizador em sítios na *web* dos municípios portugueses e, implicitamente da conformidade desses *websites* dos municípios com as normas de privacidade e segurança, uma vez que, para garantir essa conformidade, é necessária uma análise mais particular de cada município e respetivo *website,* pois existem informações que não são passíveis de recolher de forma automática, como por exemplo: as informações transmitidas ao utilizador sobre o motivo do tratamento de dados; quais os meios ao dispor para exercer o direito ao esquecimento e portabilidade, entre outros.

Nas secções seguintes serão analisados os dados recolhidos no âmbito de cada uma das temáticas abordadas nos dois capítulos anteriores (a existência de mensagens visuais que informem e/ou recolhem o consentimento para a utilização de *cookies*, a presença de "Políticas de Privacidade", a disponibilização de comunicações seguras, a existência de mecanismos de *tracking* e a eficiência dos mecanismos de defesa.

As operações de recolha de dados utilizados nesta dissertação iniciaram-se a 19 de abril de 2018, pelas 12 horas, tendo terminado a 20 de abril de 2018, às 4 horas.

5.1 Estrutura da base de dados

A estrutura da base de dados SQLite que serviu para armazenamento dos dados obtidos (que são descritos e/ou analisados nas secções seguintes) está ilustrada no Anexo B.

5.2 Informação de utilização de *cookies*

A utilização de *popups* com informação de utilização de *cookies* torna-se necessária apenas quando os *websites* utilizam *cookies*. Neste estudo, como será demonstrado no ponto 5.5, à exceção de dois municípios, todos os restantes *websites* dos municípios utilizam *cookies*. Como tal, era esperada a ocorrência deste tipo de mensagens em todos os *websites* analisados; contudo, verificou-se que apenas 26% (80 de 306) dos *websites* dos municípios portugueses apresentam estes *popups* aos seus utilizadores, conforme demonstrado no Gráfico 1.

Apesar da implementação de uma ferramenta automática para a recolha de *screenshots,* foi necessário efetuar uma análise manual e individual de cada imagem para averiguar a existência de *popus* com mensagens informativas sobre a utilização de cookies.

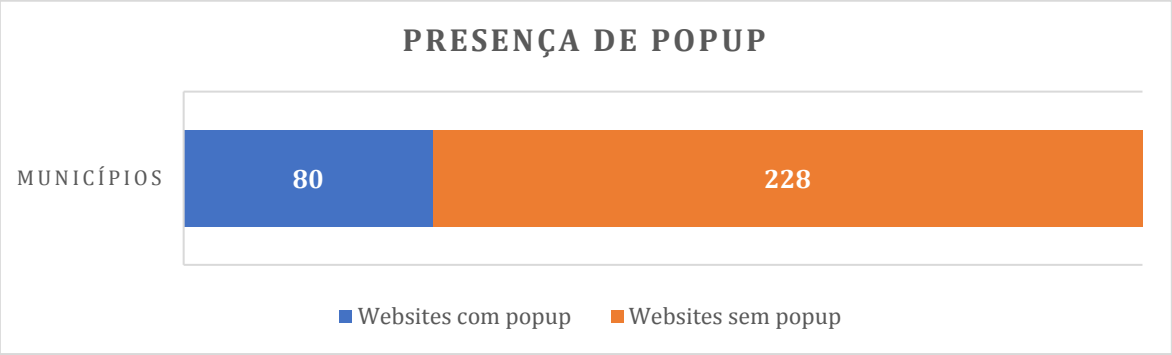


Gráfico 1 – Apresentação de informação de utilização de cookies *nos 308 websites dos municípios portugueses*

Examinando os resultados obtidos para os 80 municípios que apresentam estas mensagens e analisando o seu conteúdo criticamente, observam-se diferentes níveis no controlo de dados por parte do utilizador. Com base numa análise manual detalhada a cada *screenshot* obtido nestes municípios é possível a subdivisão nas categorias descritas na secção 3.1, conforme representa o Gráfico 2.

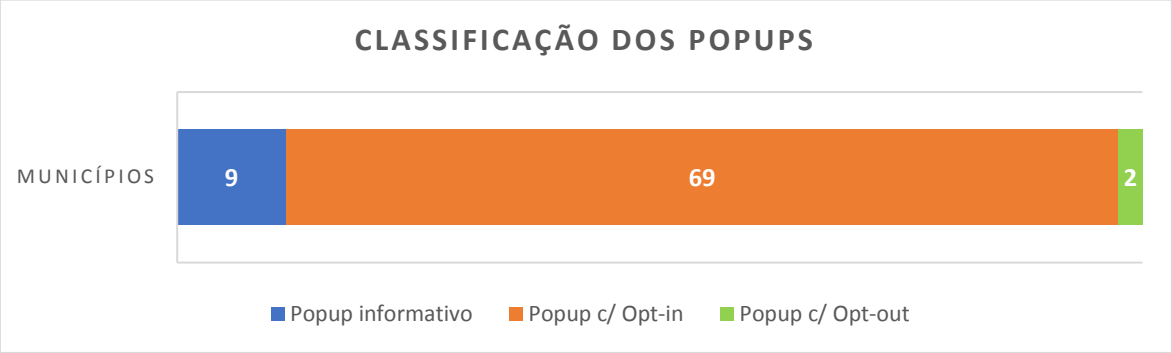


Gráfico 2 – Classificação de popups apresentados nos 80 websites dos municípios portugueses

Seria espectável que a maioria dos *websites* permitissem ao utilizador recusar a utilização de *cookies*, no entanto, apenas 2 municípios, cerca de 3% do total de *websites* com *popup*, apresentam esta opção – na Figura 11 apresenta-se o *popup* registado no *website* da Câmara Municipal da Nazaré e na Figura 12, o *popup* do Município de Albufeira.

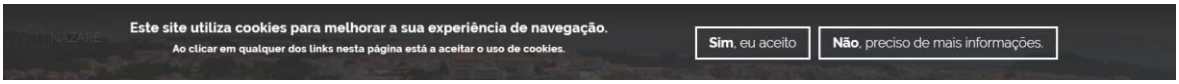


Figura 11 – Screenshot do popup com opção de opt-out do Município da Nazaré[45]

Figura 12 – Screenshot do popup com opção de opt-out do Município de Albufeira[46]

[45] http://www.cm-nazare.pt/ (screenshot a 20/04/18)
[46] http://www.cm-albufeira.pt/ (screenshot a 20/04/18)

A utilização de *popups* com *opt-out* desperta a sensibilização do utilizador, dada a necessidade clara de tomar uma opção. Acredita-se que, quando está perante uma tomada de decisão, o utilizador sentirá a necessidade de obter mais informações sobre o tema em questão.

Os *websites* que apresentam *popups* com *opt-in*, representando a maioria (cerca de 86% dos 80 referenciados), solicitam clara e explicitamente ao utilizador o consentimento para a utilização de *cookies*, i.e., existem referências a termos usualmente utilizados como forma de consentimento, tais como "Concordo" ou "Aceito", como demonstrado na Figura 13. Acredita-se que a utilização de termos ou expressões geralmente associadas a ações vinculativas, refletem uma posição concreta do utilizador perante a concordância com as condições apresentadas.

Figura 13 – Screenshot do popup com opção de opt-in do Município de Vila Nova da Barquinha[47]

Apesar da baixa representatividade, 11% dos *websites* dos municípios (9 dos 80) apresentam *popups* informativos, i.e., nos quais não é necessária qualquer ação do utilizador para consentir a utilização de *cookies*, não está refletida qualquer tomada de decisão ou compromisso direto ou explicito do utilizador. A Figura 14 exemplifica uma mensagem meramente informativa, indicando inclusivamente que ao continuar a utilizar o *website* o utilizador permite a utilização de *cookies*.

Figura 14 – Screenshot do popup informativo do Município de Arganil[48]

Se, por um lado, se pode afirmar que os municípios cujos *websites* possuem *popups* demonstram mais clareza e transparência, por outro lado, os *websites* que não

[47] http://www.cm-vnbarquinha.pt/ (screenshot a 20/04/18)
[48] http://www.cm-arganil.pt/ (screenshot a 20/04/18)

apresentam qualquer *popup* revelam uma falta de preocupação e interesse em cumprir as normas em vigor.

Apesar da existência de *popus* que permitam aceitar ou recusar a utilização de *cookies* na navegação nos seus sítios *web*, o utilizador não tem através das opções apresentadas uma verdadeira opção de escolha, i.e., independentemente da resposta do utilizador, a utilização de *cookies* ocorre e mantêm-se antes de qualquer manifestação e, mesmo, após uma recusa do consentimento.

5.3 Presença de "Políticas de Privacidade"

Sempre que, num *website*, se verifica a recolha de dados sobre os utilizadores, deve existir uma página que expresse claramente o intuito dessa recolha. Uma vez que se sabe que 306 dos 308 *websites* dos municípios utilizam mecanismos que recolhem dados, como iremos indicar na secção 5.5, esperava-se a presença de páginas com políticas de privacidade nos 306 *websites*. No entanto, só se verificou a ocorrência de *links*, com base nas palavras-chave definidas na secção 4.3, que apontam para estas políticas em cerca de 32% (98 de 308) dos *websites* dos municípios, como indicado no Gráfico 3.

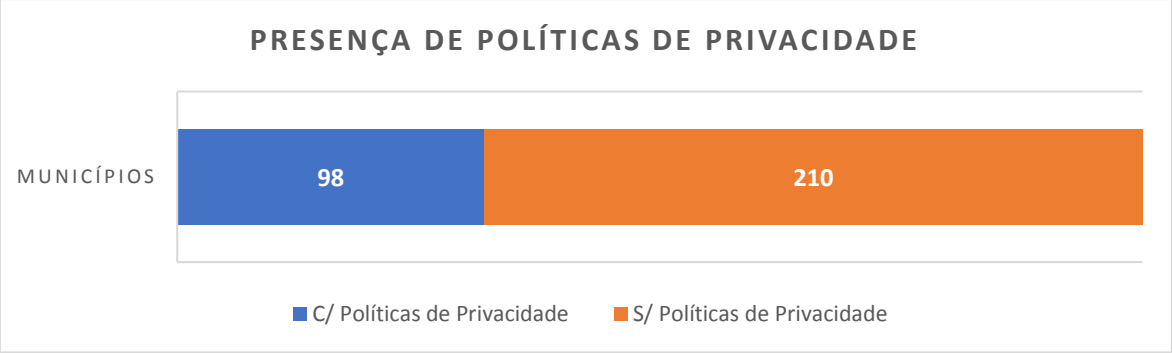


Gráfico 3 – Presença de políticas de privacidade nos 308 websites dos municípios portugueses

Para apresentar dados que permitam uma análise mais correta à conformidade com as normas, é necessário confirmar das 98 referências encontradas, quantas correspondem efetivamente a páginas válidas, i.e., cujos links apontem para uma página que contenha efetivamente informação relevante sobre as políticas de privacidade.

Através de uma análise manual aos *links* registados foi possível validar, conforme demonstrado no Gráfico 4, quais destas ocorrências:

- se referem a uma página inexistente;
- se referem a uma página cujo conteúdo não está relacionado com as políticas de privacidade;
- dizem respeito a páginas que contenham efetivamente informações sobre a recolha e tratamento de dados.

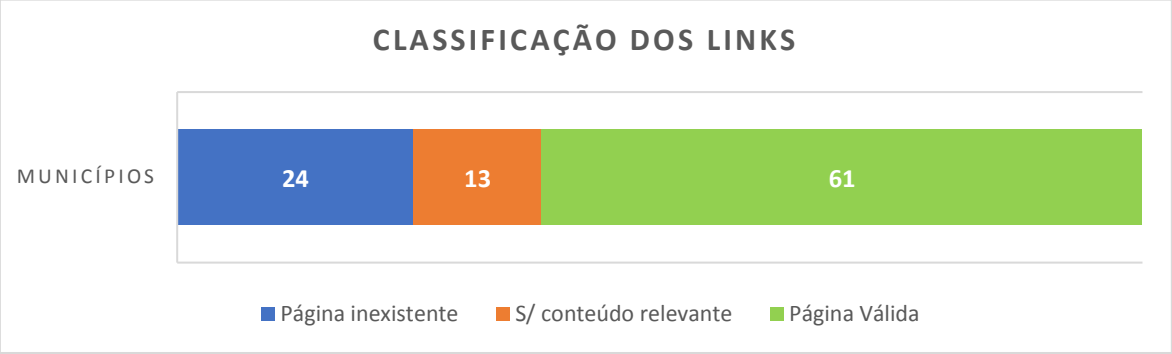


Gráfico 4 – Classificação dos links obtidos relativos a Políticas de Privacidade nos 98 municípios portugueses

Foram encontrados 24 *websites* que apresentavam *links*, cujos nomes incluíam as expressões definidas e que continham redirecionamentos para a mesma página visitada. Estes redirecionamentos são possíveis através da utilização do caracter # que, quando contido na referência do *link*, permite a realização de um deslocamento para um objeto contido no *Document Object Model* (DOM) da mesma página visitada. Nestas situações não existe sequer uma página ou zona especifica para as políticas de privacidade, onde o utilizador se possa informar sobre a recolha e tratamento de dados.

Na figura 15 é demonstrada a ocorrência deste tipo de *links*, sendo verificado que a página efetivamente não existe.

Figura 15 – Screenshot efetuado ao link registado para a política de privacidade do Município de Aveiro[49]

Em 13 dos casos verificados, ao contrário do exemplo anterior, os *links* registados dizem efetivamente respeito a outras páginas existentes, mas que não apresentam qualquer conteúdo válido, como exemplificados nas Figuras 16 e 17.

O *website* do Município de Sesimbra (Figura 16) apresenta uma página sem qualquer texto sob o título "Política de Privacidade".

O *website* do Município de Santos Tirso (Figura 17) apresenta uma página intitulada "Política de privacidade e segurança", apenas indicando de seguida que o conteúdo estará "brevemente disponível". Apesar de ser demonstrada alguma preocupação com a proteção de dados, já que existe um espaço dedicado a este tema, considera-se que estes casos são, no âmbito da análise à conformidade, equivalentes aos casos categorizados como "páginas sem conteúdo relevante".

[49] http://www.cm-aveiro.pt/www/# (screenshot a 20/04/18)

Figura 16 – Screenshot da página de políticas de privacidade da CM Sesimbra[50]

Figura 17 – Screenshot da página de políticas de privacidade da CM Santo Tirso[51]

A maioria dos *links* encontrados nos *websites* dos municípios, cerca de 62% (61 dos 98), apontam efetivamente para páginas que contêm informações sobre as políticas de privacidade, como exemplificado na Figura 18.

[50] http://www.cm-sesimbra.pt/pages/1080 (screenshot a 20/04/18)
[51] https://www.cm-stirso.pt/frontoffice/pages/412 (screenshot a 20/04/18)

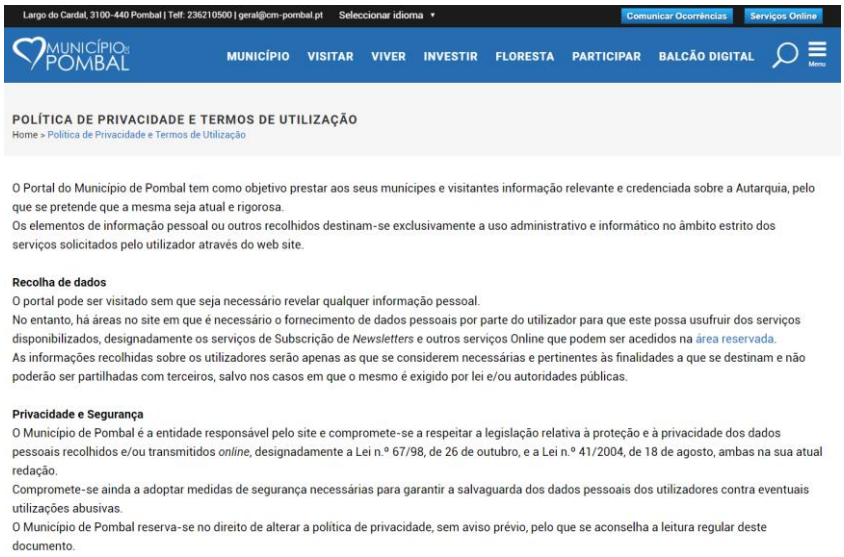


Figura 18 - Screenshot de excerto da página de políticas de privacidade da CM Pombal[52]

Apesar de estas páginas existirem e apresentarem conteúdo identificável como estando contextualizado no tema, este conteúdo não representa uma garantia de conformidade com as normas de privacidade e segurança, uma vez que seria necessário analisar manual e profundamente cada uma dessas páginas e respetivo conteúdo.

Não obstante, numa análise superficial efetuada para validar se efetivamente existia conteúdo relevante nestes *links*, verificou-se o exemplo anterior, como um bom exemplo das informações prestar, uma vez que o Município de Pombal terá dedicado um esforço para informar os dados que são recolhidos, a finalidade da recolha desses mesmos dados, informa a possibilidade de retificar esses dados, informa sobre a utilização de *cookies* e o propósito dessa utilização.

5.4 Disponibilização de Comunicações Seguras

A análise efetuada permite averiguar a segurança dos dados em transferência na navegação em *websites* dos municípios, tendo-se verificado que apenas cerca de

[52] https://www.cm-pombal.pt/politica-de-privacidade-e-termos-de-utilizacao/ (screenshot a 20/04/18)

24% (75) dos 308 municípios disponibilizam protocolos de comunicações seguras, conforme apresentado no gráfico 5.

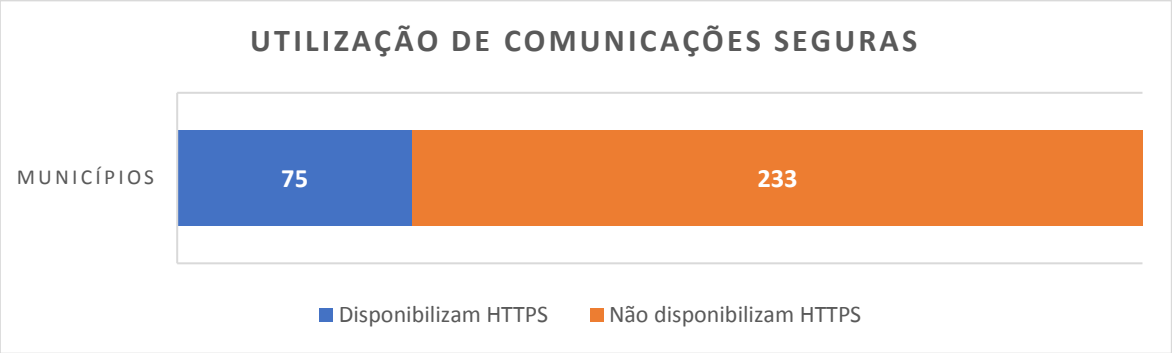


Gráfico 5 – Utilização de protocolos de comunicações seguras nos websites dos 308 municípios portugueses

Não disponibilizando protocolos de comunicações seguras aos utilizadores dos seus *websites*, a maioria dos municípios não podem garantir a segurança dos dados transmitidos pelo utilizador na navegação *web*.

Após efetuar esta validação, analisou-se em cada um dos 75 *websites* que disponibilizam o protocolo HTTPS, em quantos destes acontece o redirecionamento para HTTPS quando é solicitado um pedido HTTP, conforme demonstrado no Gráfico 6.

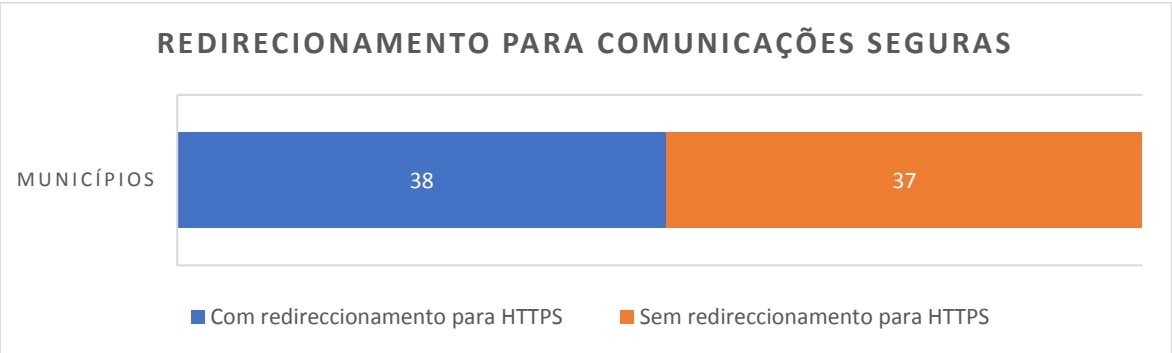


Gráfico 6 – Redirecionamento para protocolos de comunicações seguras nos websites dos 75 municípios que disponibilizam estes protocolos

Tendo em conta que os municípios que disponibilizam o protocolo HTTPS representam uma baixa percentagem sobre o total de municípios, seria de esperar que

a maioria desse prioridade à utilização deste protocolo redirecionando o tráfego mesmo que fosse solicitado o pedido HTTP. No entanto, apenas metade dos municípios prioriza o protocolo HTTPS, sendo que os restantes continuam a não poder garantir total segurança dos dados em trânsito.

5.5 Existência de Mecanismos de *Tracking*

Como referido anteriormente, a análise à existência de mecanismos de *tracking* foi efetuada em dois procedimentos distintos: uma análise individual a cada conjunto de páginas dos municípios e uma análise coletiva e consecutiva aos vários *websites* dos municípios.

Em ambas as análises se recolheram informações relativas à utilização de *cookies*, a comunicações HTTP a domínios distintos do visitado, os acessos à API do *browser* e a existência de LSOs. Considera-se que através da análise crítica a este tipo de informação é possível identificar as técnicas de *tracking* mais comuns.

Na análise individual à utilização de *cookies* nos sítios *web* dos municípios, foi possível verificar, à data do estudo, que apenas 2 dos 308 municípios portugueses não colocaram *cookies* (Município de Pinhel[53] e Município de Porto Santo[54]), como demonstrado no Gráfico 7.

Gráfico 7 – Utilização de cookies nos 308 websites dos municípios

[53] http://cm-pinhel.pt/ (acedido em 21/04/2018)
[54] http://cm-portosanto.pt/ (acedido em 21/04/2018)

Importa analisar nos 306 *websites* que colocam *cookies*, se estas pertencem ao domínio do respetivo município, sendo possível ser acedido apenas por esse domínio, ou se, por outro lado, pertencem a um domínio diferente do município, podendo ser acedidas em visitas a páginas de domínios diferentes do município.

Observou-se que 227 dos 306 municípios que colocam *cookies*, colocam *cookies* de domínios diferentes do seu – *third party cookies* – como demonstrado no Gráfico 8.

Gráfico 8 – Utilização de third party cookies nos websites dos 306 municípios que utilizam cookies

Considerou-se que a utilização de *third party cookies* por parte dos municípios não é uma prática transparente – como se verificou nos pontos 5.2 e 5.3, na maioria dos casos não existe qualquer informação sobre a recolha e processamento dos dados (seja à entrada do *website*, seja numa secção especifica para o efeito). A existência de *cookies* de terceiros aliada à falta ou a fraca comunicação da utilização destes mecanismos representa uma falta de preocupação com o cumprimento das normas.

Dada a diversidade de objetivos ou propósitos para a utilização de *cookies*, considerou-se pertinente avaliar a percentagem de *third party cookies* face ao número total de *cookies* diferentes utilizadas por município. Os resultados obtidos e demonstrados no Gráfico 9, permitem afirmar que 147 dos 227 municípios que utilizam *third party cookies*, possuem mais *cookies* de outros domínios do que do próprio domínio.

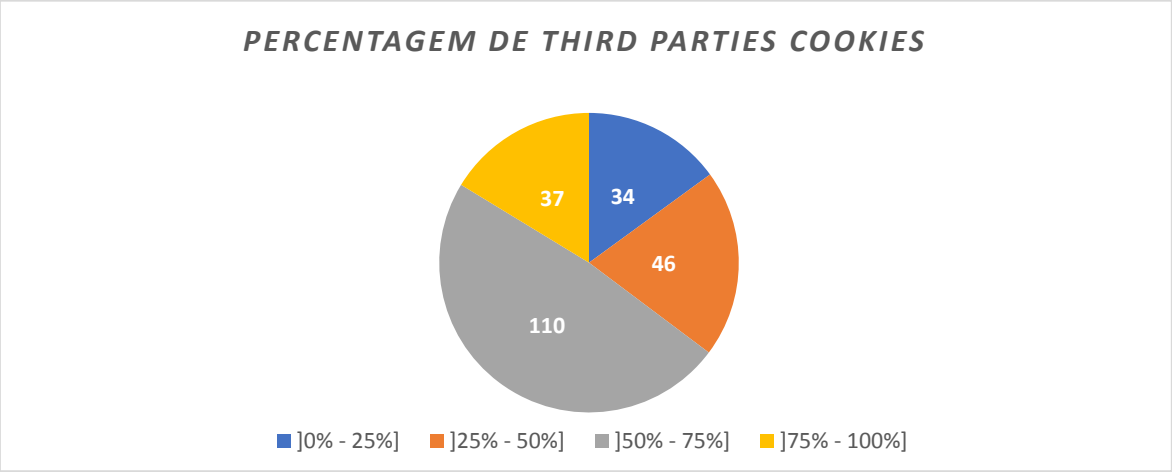


Gráfico 9 – Percentagem de third parties cookies em relação ao número total de cookies utilizadas

Nas ocorrências onde se verifica uma elevada percentagem de *third party cookies*, é percetível que as finalidades para a utilização de *cookies* não são, exclusivamente, o controlo de sessão, a otimização do *website* e a oferta de conteúdo personalizado. Desta forma, para garantir que se está, efetivamente, perante um mecanismo de *tracking*, considerou-se importante verificar quais os domínios das *third party cookies* mais utilizadas. No Gráfico 10, estão representados os 9 domínios de terceiros mais utilizados pelos municípios.

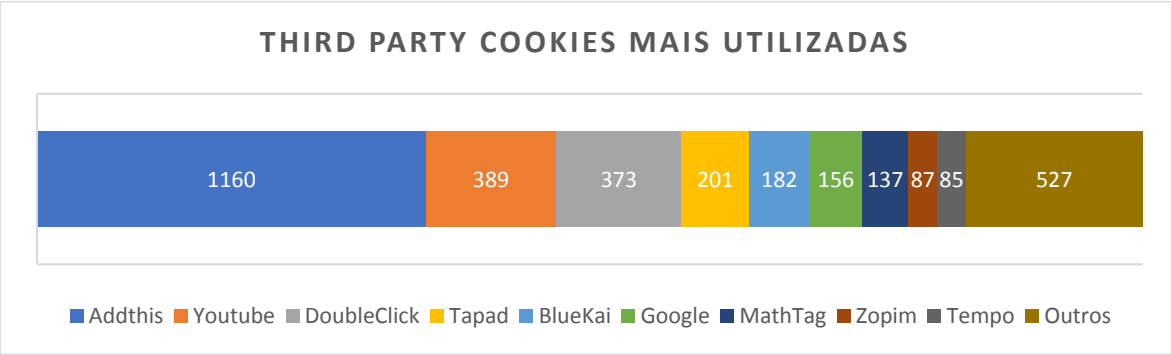


Gráfico 10 – Tirth party cookies mais utilizadas nos 306 websites dos municípios portugueses que utilizam cookies

É claro que a utilização de determinadas *third parties cookies* não representa *tracking*, enquanto recolha de dados pessoais do utilizador: o *Google Analytics*[55], por

[55] https://www.google.com/analytics/

exemplo, que é comummente utilizado para fins estatísticos, o DoubleClick[56] que permite monitorizar interesse em anúncios [40], ou o Tempo[57], que permite prestar informações úteis ao utilizador. O Youtube[58] e o Zopim[59] são exemplos de *third parties cookies* que existem por terem sido incluídas nos *websites* dos municípios funcionalidades especificas, respetivamente vídeos incorporados e ferramentas de chat. Por outro lado, a utilização de *third parties cookies*, como o Addthis[60], o Tapad[61], o BlueKai[62] e o MathTag[63], tem como finalidade a criação de perfis de navegação [41].

O registo das comunicações HTTP, como demonstrado nos Gráficos 11 e 12, permitiu observar que na maioria dos *websites* analisados, a existência de *scripts* incluídos nas páginas HTML que causam um pedido HTTP *request* a um domínio diferente do município, não pode ser considerada como mecanismo de *tracking,* uma vez que estes são utilizados apenas para embelezar ou otimizar a página HTML – desta forma não é recolhida qualquer informação do utilizador.

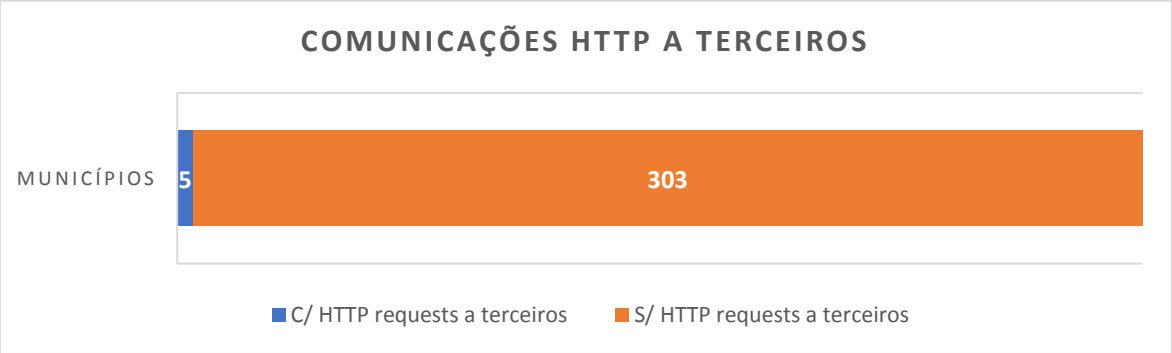


Gráfico 11 – Municípios portugueses nos quais se registaram comunicações HTTP a domínios diferentes do respetivo website

Para garantir uma correta análise ao gráfico anterior, a informação deve ser complementada com o *ratio* entre a quantidade de comunicações HTTP a terceiros e o

[56] https://www.doubleclickbygoogle.com/
[57] https://tempo.pt/
[58] https://www.youtube.com/
[59] https://pt.zopim.com/
[60] http://www.addthis.com/
[61] https://www.tapad.com/privacy-policy (acedido em 22/04/2018)
[62] https://www.oracle.com/corporate/acquisitions/bluekai/index.html
[63] http://www.mediamath.com/privacy-policy/

total das comunicações HTTP registadas por município, como ilustrado no Gráfico 10. Assim, garante-se uma visão mais detalhada sobre a necessidade deste tipo de pedidos.

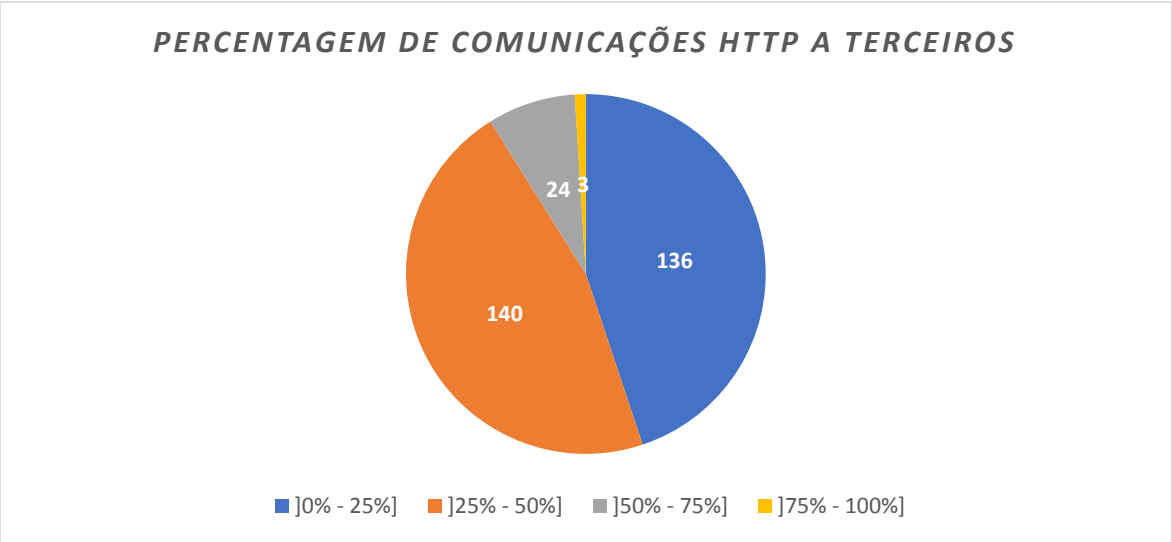


Gráfico 12 - Percentagem de comunicações HTTP a terceiros vs número total de pedidos

Neste gráfico é percetível que a grande maioria dos *websites* dos municípios apresentam mais comunicações HTTP ao seu próprio domínio do que a terceiros, no entanto, existem 27 municípios que apresentam mais comunicações HTTP a terceiros do que ao seu próprio domínio. Destes 27, os municípios de Fafe, Vila do Bispo e Fundão destacam-se pela grande percentagem de comunicações HTTP a terceiros, com mais de 75% de comunicações HTTP a terceiros. Esta desproporcionalidade indica uma grande falta de preocupação com a privacidade dos dados dos seus utilizadores, uma vez que existem mais dados partilhados com terceiros do que com o próprio município visitado.

As comunicações HTTP a terceiros podem não representar, por si só, um mecanismo de *tracking*, tal como na utilização de *third party cookies*. No entanto, os domínios registados, conforme ilustrado no gráfico 13, demonstram a partilha de dados e a possibilidade de *tracking*, uma vez que apesar destas comunicações permitirem aos municípios a análise estatística ao *website* (Google) ou a inclusão de *widgets* do Facebook ou vídeos do Youtube, as respetivas entidades terceiras recolhem e tratam os dados sem intervenção dos municípios. Também no caso do Addthis, a

partilha de dados com terceiros pode ocorrer para fins publicitários e comerciais, originando a segmentação de utilizadores.

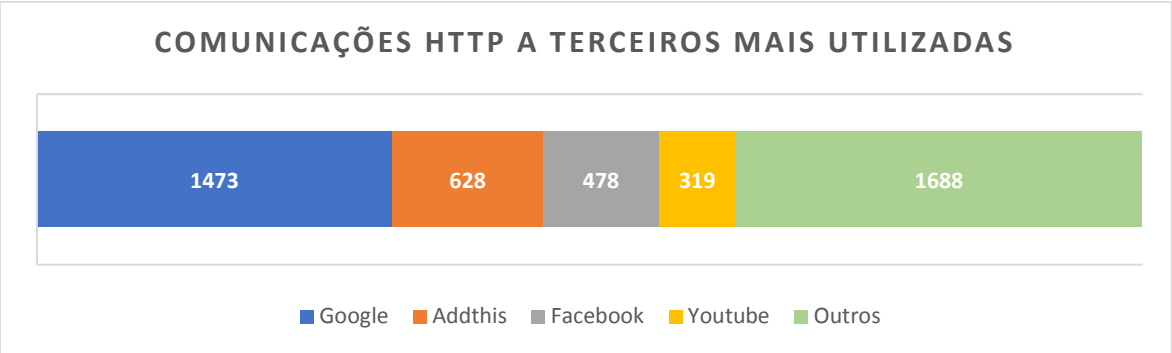


Gráfico 13 – Domínios das comunicações HTTP a terceiros mais utilizados nos websites dos municípios portugueses

Na análise aos sítios *web* dos municípios pretendia-se também registar a utilização de *flash cookies*, no entanto, em nenhuma das páginas visitadas dos 308 municípios foram encontradas ocorrências de LSOs. Estes resultados podem ser justificados pelo facto de o *browser* Mozilla Firefox não ter como predefinição a autorização do *plugin*, não obstante, considera-se que esta predefinição que geralmente existe nos *browser*es representa um alerta de segurança ao utilizador, como tal, a sua aceitação representaria uma forma de consentimento, ainda que de forma não explicita.

Confirmou-se, nesta análise individual, que todos os municípios utilizam *Javascript calls,* como seria espectável até para que se permita a otimização e a correta apresentação do *site*. No entanto, uma recolha exaustiva de dados através destes mecanismos e da sua correlação, permite reconhecer, com um elevado grau de certeza, o terminal de navegação, podendo, assim, identificar o padrão de navegação de um utilizador. No Gráfico 14 são demonstradas as API *calls* mais encontradas nesta análise.

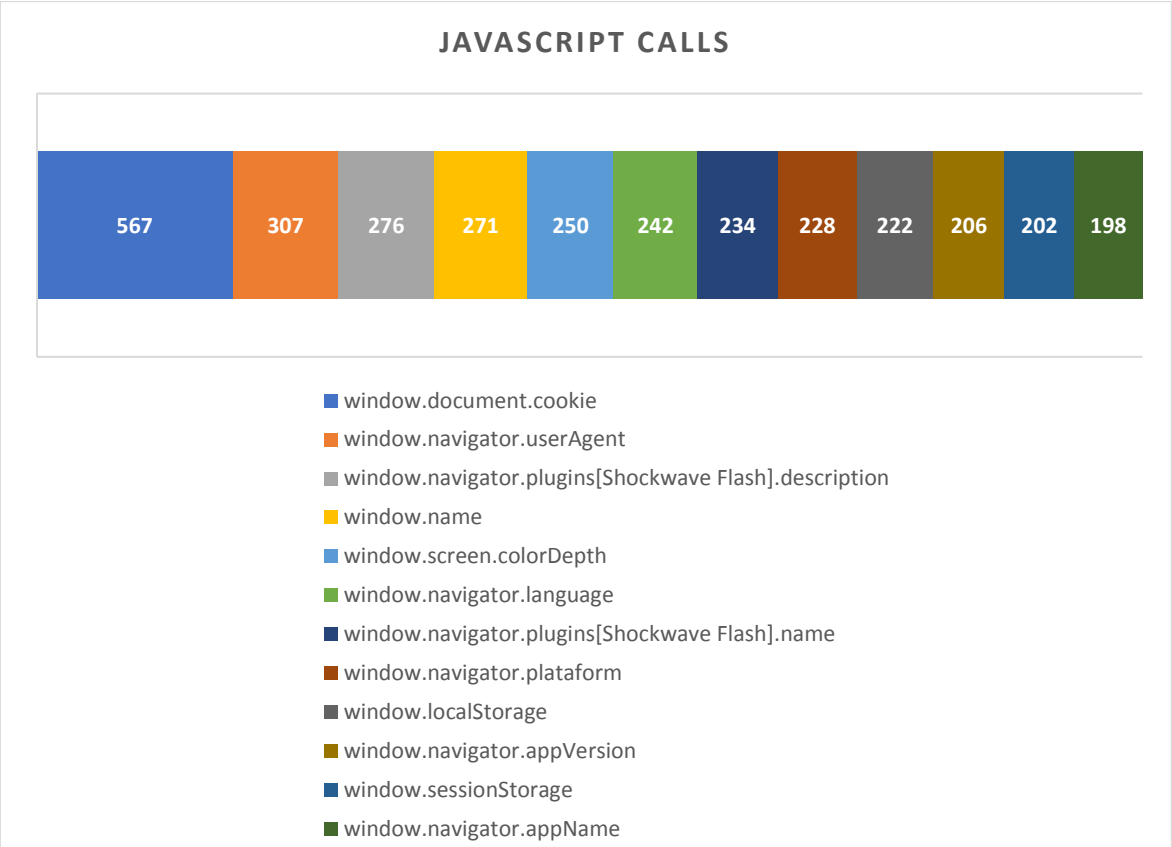


Gráfico 14 – Javascript calls utilizados nos websites dos municípios portugueses

Tendo em conta os resultados recolhidos numa análise independente a cada conjunto de páginas de cada município, utilizou-se uma mesma instância do *browser* para visitar um grande conjunto de páginas dos vários municípios, de forma a perceber que um único *tracker* pode seguir a navegação de um utilizador entre vários *municípios*. Os resultados obtidos permitem observar que diversas *cookies* são criadas numa visita a uma página *web* de um município e são acedidas e/ou alteradas em páginas de outros municípios que foram posteriormente visitadas.

Na Figura 19 está demonstrado um dos exemplos encontrados acesso e alteração de uma mesma *cookie* em municípios.

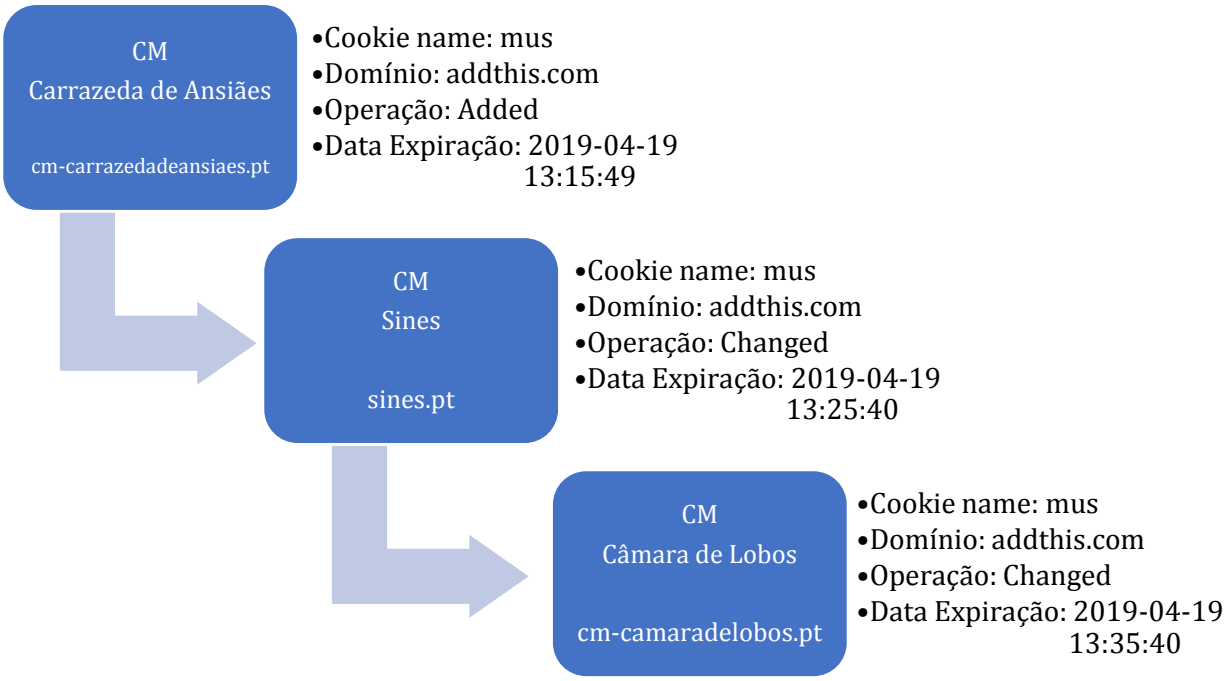


Figura 19 – Esquema de parte da utilização de uma cookie na navegação em diferentes municípios

Neste exemplo, uma *cookie* do Addthis foi criada na visita ao *website* da Câmara Municipal de Carrazeda de Ansiães, tendo sido registada uma data de expiração; posteriormente, na visita ao *website da* Câmara Municipal de Sines, tendo sido acedida a mesma *cookie*, alterando apenas a data de expiração; uma alteração semelhante foi registada aquando da visita à Câmara Municipal de Câmara de Lobos, sendo novamente aumentada a data de expiração.

Através de uma análise manual aos dados recolhidos, que estão registados numa base de dados SQLite, foi possível perceber que determinadas *cookies* foram acedidas por diferentes domínios. Conforme enumerado no Anexo C, identificaram-se as seguintes *cookies* :

- Addthis – as *cookies* deste domínio foram acedidas por 78 municípios;
- Farmácias de Serviços[64] – as *cookies* deste domínio foram acedidas por 6 municípios;
- Issuu[65] – as *cookies* deste domínio foram acedidas por 5 municípios;
- Tempo – as *cookies* deste domínio foram acedidas por 14 municípios;

[64] http://www.farmaciasdeservico.net/
[65] https://issuu.com/

- Youtube – as *cookies* deste domínio foram acedidas por 72 municípios;
- Zopim – as *cookies* deste domínio foram acedidas por 19 municípios.

5.6 Eficiência dos Mecanismos de Defesa

É importante perceber até que ponto os mecanismos de defesa disponíveis cumprem efetivamente a sua função e protegem o utilizador e os seus dados pessoais. Para garantir uma correta análise desta eficiência, efetuaram-se dois pares de análises à navegação em *websites* de diferentes municípios, tendo sido analisadas 15 páginas de cada um dos 10 municípios selecionados (Portimão, Salvaterra de Magos, Lajes das Flores, Arouca, Carrazeda de Ansiães, Resende, Vila Nova de Paiva, Sines, Mesão Frio e Aveiro), com base nos resultados recolhidos anteriormente, i.e., foram escolhidos os municípios que demonstraram presença de *third party cookies* e *third party requests*.

Para verificar a eficiência do DNT, numa primeira análise, sem a utilização do mecanismo de defesa, recolheram-se os resultados demonstrados no Gráfico 15.

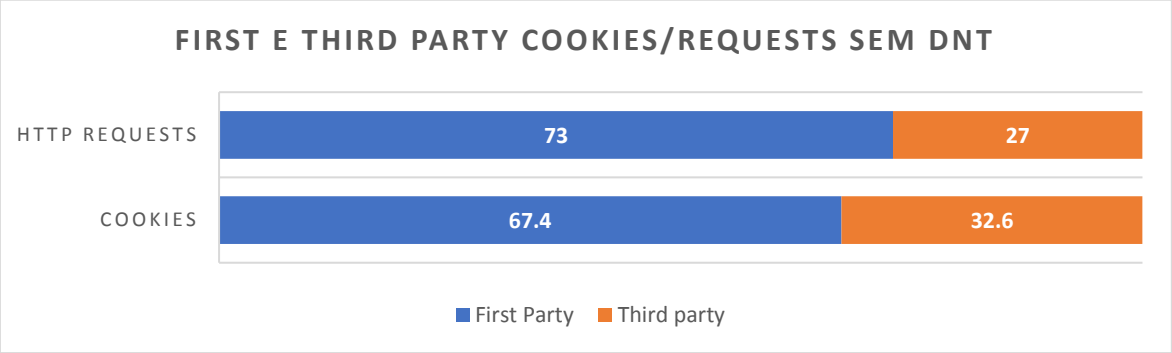


Gráfico 15 – Percentagem de first e third party cookies e HTTP requests sem DNT

Estes dados, que mostram claramente a elevada quantidade de *third party cookies* e comunicações HTTP a terceiros nas 150 páginas visitadas face aos totais, funcionarão como referência para os valores recolhidos numa navegação idêntica à anterior, porém utilizando o DNT conforme descrito na secção 4.6. Os resultados registados nesta navegação com DNT estão apresentados no Gráfico 16.

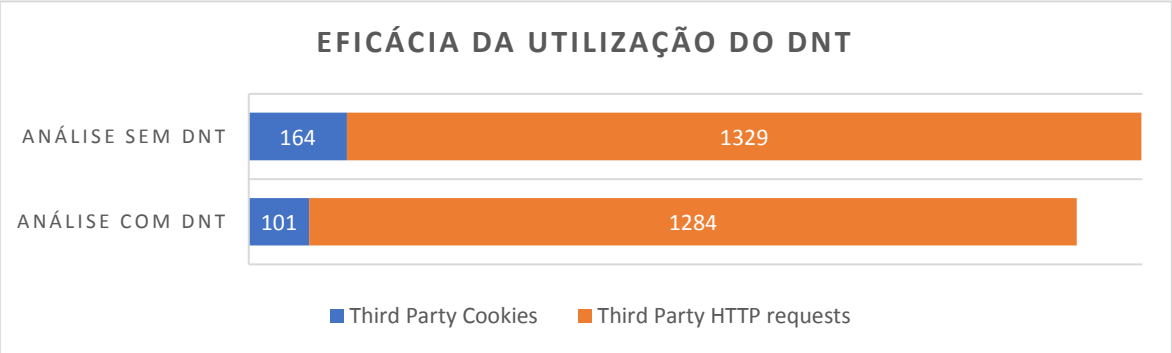


Gráfico 16 – Comparação de resultados nas análises à eficácia na redução de third party cookies e HTTP requests com a utilização do DNT

Como se pode verificar na comparação dos dois gráficos anteriores a utilização do DNT na navegação, resultou numa ligeira diminuição do número de *third party cookies* – da primeira para a segunda análise registou-se uma diminuição de 63 *cookies* de terceiros, uma variação de cerca de 9% face ao número total de *cookies*. Esta diferença é muito menor no caso dos HTTP *requests* a terceiros, que apresentam uma variação de apenas 1% (cerca de 40 registos).

Apesar de ser expectável que este mecanismo de defesa tivesse uma eficiência superior dado ser uma implementação protocolar, nesta análise percebemos que o DNT não apresenta resultados significativos, uma vez que como referido na secção 2.6.4, os servidores não são obrigados a honrar o DNT.

Na primeira análise, como análise-padrão, efetuada para testar a eficiência do Ghostery, obtiveram-se os resultados demonstrados no Gráfico 17.

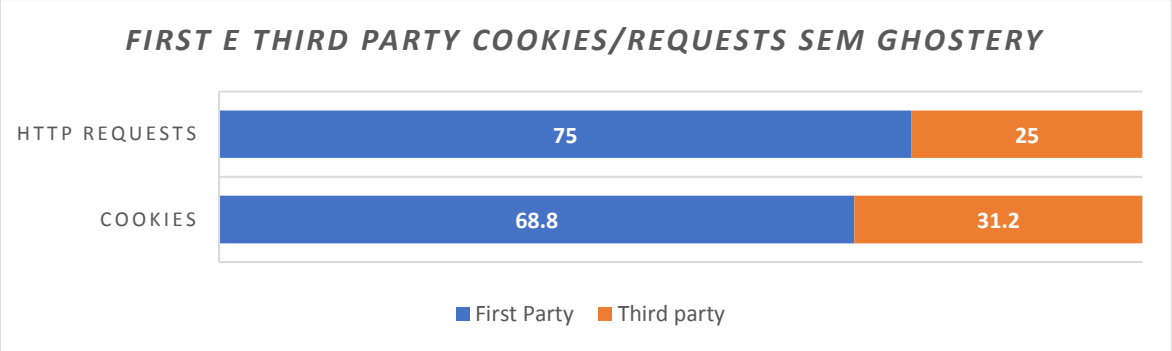


Gráfico 17 – Percentagem de first e third party cookies e HTTP requests sem Ghostery

Estes resultados, que comprovam a necessidade de efetuar análises padrão independentes para cada um dos mecanismos, traduzem em percentagem os *ratios first vs third party* das 536 *cookies* impostas e dos 5075 HTTP *requests* registados. Após esta análise, procedeu-se a uma nova análise, desta vez utilizando o Ghostery e obtendo os resultados expostos no Gráfico 18.

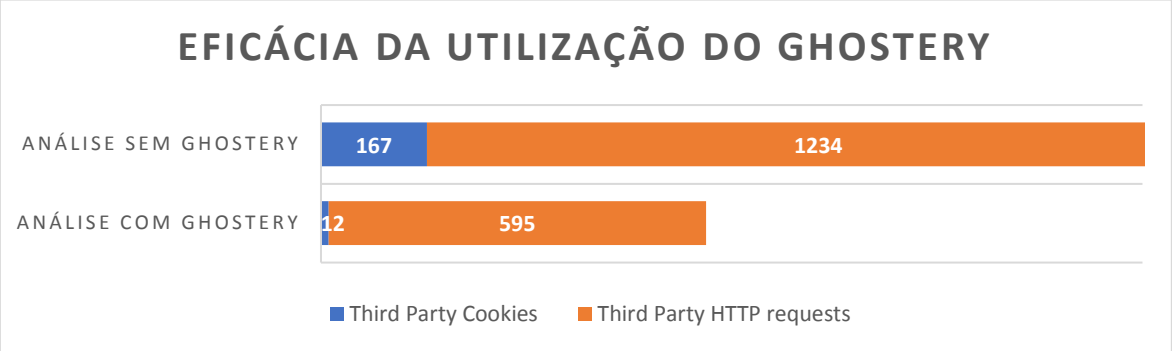


Gráfico 18 – Comparação de resultados nas análises à eficácia na redução de third party cookies e HTTP requests com a utilização do Ghostery

Das análises efetuadas para a averiguação da eficiência do Ghostery, podemos afirmar que este mecanismo de defesa demonstra resultados postivos: verificou-se que o Ghostery bloqueou mais de 50% das comunicações HTTP a terceiros e impediu que 93% das *third party cookies* fossem impostas. Verificou-se também que os 7% de *third party cookies* que nesta análise foram impostas poderiam também ser bloqueadas se se tivesse optado por essa configuração, tal como referido na secção 4.6. As *third party cookies* que foram permitidas nesta análise pertencem aos domínios Google.com e Youtube.com.

Observando os resultados recolhidos e apresentados nos gráficos 15 a 19, pode concluir-se que a utilização do Ghostery representa um efetivo mecanismo de defesa que o utilizador pode configurar para limitar a recolha dos seus dados pessoais e, assim, proteger a sua privacidade.

6 Conclusões

Neste último capítulo, em duas secções, serão feitas considerações sobre os resultados obtidos, sobre a contribuição desta dissertação para os seus leitores e sugestões de trabalho e análises futuras.

6.1 Considerações Finais

Da análise aos resultados desmontados na secção 5.5, conclui-se que a maioria dos *websites* dos municípios recolhem ativamente informação sobre o utilizador pelo que era esperado que informassem o utilizador dessa recolha e tratamento de dados quer através de um *popup* quer disponibilizando uma página com as políticas de privacidade. No entanto, 74% dos municípios não apresenta nenhum *popup* à entrada do *website* e os que apresentam não cumprem necessariamente os requisitos legais, não disponibilizando opções claras ao utilizador. Por outro lado, só cerca de 20% dos *websites* dos municípios têm uma secção dedicada à comunicação da política de

privacidade. Ainda assim, como referido anteriormente, não foi efetuada uma análise ao conteúdo de cada uma dessas páginas ou secções, pelo que apesar de se considerar que estes 61 municípios demonstram alguma preocupação e respeito pela privacidade de dados, não se pode afirmar que estes estão em total conformidade as normas em vigor.

Tendo em conta que os municípios recorrem vastas vezes à troca de informação com os utilizadores dos seus *websites* através de, por exemplo, submissão de documentos, preenchimento de formulários, era esperado que a maioria utilizasse um canal de comunicação segura. Apesar disso, constatou-se que apenas 75 dos 308 websites utilizam o protocolo HTTPS e destes apenas 50% prioriza o canal de comunicação segura.

Apesar de existirem 227 câmaras municipais que utilizam *third party cookies* nos seus *websites,* cuja utilização pode ser considerada normal, por exemplo, para a otimização estatística do *website* ou personalização de conteúdo, como a meteorologia, não se pode afirmar que os *websites* demonstram preocupação com a privacidade dos seus utilizadores, uma vez que como referido anteriormente, na maioria dos casos, não há nem uma comunicação explicita, nem um pedido de consentimento do utilizador.

Mesmo tendo em consideração, como já referido, que as comunicações HTTP a terceiros não representam por si só um mecanismo de *tracking*, deve destacar-se que cerca 27% dos *websites* dos municípios apresenta uma taxa de pedidos a terceiros superior a 50% face ao número total de comunicações HTTP. Deste facto depreende-se que não existe preocupação com a inclusão de *scripts* alojados em servidores de terceiros que podem representar uma ameaça à privacidade do utilizador, se e quando as intenções de quem desenvolveu o *script* não são as mais corretas.

Na análise conjunta aos vários *websites* dos municípios foi observada a presença de vários mecanismos de *tracking*, através da inclusão de *scripts* de terceiros em vários *websites*, possibilitando o estudo dos interesses e dos comportamentos do utilizador.

Na análise à eficiência dos mecanismos de defesa estudou-se o comportamento do DNT e do Ghostery. Este primeiro mecanismo demonstrou claramente resultados opostos ao esperado: impediu um baixo número de *third party cookies* e de HTTP *requests* na navegação em vários *websites* de municípios. Por outro lado, o Ghostery

apresentou uma grande taxa de sucesso, tendo impedido a totalidade das *third party cookies*, à exceção dos domínios Google.com e Youtube.com (que estão permitidos nos filtros padrão do Ghostery.) Assim, pode afirmar-se que o Ghostery é um mecanismo de defesa eficiente na navegação *web*, cumprindo a função de proteger a privacidade dos dados.

6.2 Trabalhos Futuros

Observando o trabalho desenvolvido e os resultados recolhidos considera-se importante que no futuro se considere:

- realizar novas análises aos *websites* dos municípios em períodos definidos, como 6 e 12 meses após a entrada em vigor do RGPD. Tendo em conta que a recolha de dados ocorreu antes da entrada em vigor das normas europeias e que os municípios podem levar algum tempo a adaptar-se a estas normas, considera-se de extrema relevância uma análise à evolução da adaptação dos *websites* desta área homogénea;
- realizar uma análise à privacidade do utilizador em sítios na *web* de uma outra área homogénea – dada a disparidade de interesses na presença *web* das diversas áreas;
- desenvolver e disponibilizar uma aplicação *web* que possibilite o estudo da privacidade num domínio a definir pelo utilizador. A ferramenta a desenvolver deve também gerar um relatório detalhado com base nas informações recolhidas.

Bibliografia

[1] T. Berners-Lee and M. Fischetti, *Weaving the Web : the original design and ultimate destiny of the World Wide Web by its inventor*. HarperSanFrancisco, 1999.

[2] A. Thierer, J. Brito, L. Downes, A. Marcus, and R. Radia, "The Pursuit Of Privacy In A World Where Information Control Is Failing," *Harv. J. Law Public Policy*, vol. 36, 2013.

[3] C. Barnard and S. Peers, *European Union Law (second edition)*. Oxford University Press, 2017.

[4] Assembleia da República (VII revisão constitucional), "Constituição da República Portuguesa," 2005.

[5] L. De Almendra and F. Pires, "DIREITO À PRIVACIDADE NO ÂMBITO DA SOCIEDADE DA INFORMAÇÃO: reflexões em torno da questão nos inícios do século XXI," Mestrado Científico, Faculdade de Direito da Universidade de Coimbra, 2014.

[6] ACEPI and IDC, "Estudo Anual da Economia e da Sociedade Digital em Portugal," 2016. [Online]. Available: https://acepi.pt/download.php?f=ACEPI - Estudo Economia Digital 2016 - Resumo.pdf.

[7] Gávea and Universidade do Minho, "Presença na Internet das Câmaras Municipais Portuguesas em 2016 Estudo sobre Local e-Government em Portugal," 2017. [Online]. Available: http://gavea.dsi.uminho.pt/wp-content/uploads/2017/05/Ipic2016_Pub2017_VFINAL.pdf. [Accessed: 02-Jun-2018].

[8] Observatório da Sociedade da Informação, "Presença na Internet das Câmaras Municipais Portuguesas em 2016, Estudo sobre Local e-Government em Portugal," 2017.

[9] Assembleia da República, "Lei nº67/98 de 26 de Outubro de 1998," *Diário da República - I Série-A*, pp. 5536–5546, 1998.

[10] Assembleia da República, "Lei nº46/12 de 29 de agosto de 2012," *Diário da República - I Série-A*, 2012.

[11] Parlamento Europeu and Conselho Europeu, "Regulamento (UE) 2010/87 do Parlamento Europeu e do Concelho de 5 de fevereiro de 2010," 2010.

[12] N. Robinson, H. Graux, M. Botterman, and L. Valeri, "Review of the European Data Protection Directive," *RAND Eur.*, 2009.

[13] Parlamento Europeu, "Proteção da privacidade num mundo interligado Um quadro europeu de proteção de dados para o século XXI," *COM*, vol. 25, no. 9, 2012.

[14] Parlamento Europeu and Conselho Europeu, “Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho de 27 de abril de 2016,” 2016.

[15] Allen & Overy, “The EU General Data Protection Regulation,” 2017. [Online]. Available: http://www.allenovery.com/SiteCollectionDocuments/Radical changes to European data protection legislation.pdf.

[16] M. Wnuk and R. Atterer, “Knowing the user’s every move: user activity tracking for website usability evaluation and implicit interaction,” *15th Int. World Wide Web Conf.*, 2006.

[17] M. Leonardi, “Tutela e privacidade na Internet,” 2012. [Online]. Available: http://leonardi.adv.br/wp-content/uploads/2012/01/mltpi.pdf.

[18] D. J. Solove, “GDPR Whiteboard,” 2017.

[19] Grupo de Trabalho do Art. 29ª para a Proteção de Dados and CNPD, “Orientações relativas à Avaliação de Impacto sobre a Proteção de Dados (AIPD) e que determinam se o tratamento é «suscetível de resultar num elevado risco» para efeitos do Regulamento (UE) 2016/679,” 2017.

[20] B. Krishnamurthy and C. E. Wills, “On the leakage of personally identifiable information via online social networks,” *Proc. 2nd ACM Work. Online Soc. networks*, 2009.

[21] J. Mayer, “Tracking the trackers: To catch a history thief,” 2011. [Online]. Available: http://cyberlaw.stanford.edu/blog/2011/07/tracking-trackers-catch-history-thief.

[22] David Heitmeyer, “HTTP Cookies,” *Harvard Grab Bag*, 2011. [Online]. Available: https://cscie12.dce.harvard.edu/lecture_notes/2011/20110504/slide3.html. [Accessed: 16-Apr-2018].

[23] A. Soltani, S. Canty, and C. Hoofnagle, “Flash cookies and privacy,” *SSRN*, 2009.

[24] D. Herrmann, R. Wendolsky, and H. Federrath, “Website fingerprinting: attacking popular privacy enhancing technologies with the multinomial naïve-bayes classifier,” in *Proceedings of the 2009 ACM workshop on Cloud computing security - CCSW ’09*, 2009, p. 31.

[25] “Postfix Virtual.” [Online]. Available: http://www.postfixvirtual.net/. [Accessed: 28-Feb-2018].

[26] J. Schmidt, “Das Like-Problem,” *Heise Secur.*, 2011.

[27] “ECMAScript - Language Specification,” 2015. [Online]. Available: https://www.ecma-international.org/ecma-262/5.1/.

[28] E. Phetteplace and M. K. Kern, “Hardening the Browser - Protecting Patron Privacy on the Internet,” 2012. [Online]. Available: http://eprints.rclis.org/16837/1/Hardening-the-Browser.pdf.

[29] G. Acar *et al.*, “FPDetective: Dusting the Web for Fingerprinters,” 2013. [Online]. Available:

https://www.esat.kuleuven.be/cosic/publications/article-2334.pdf.

[30] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz, "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild," 2014. [Online]. Available: https://securehomes.esat.kuleuven.be/~gacar/persistent/the_web_never_forgets.pdf.

[31] S. Englehardt, C. Eubank, P. Zimmerman, D. Reisman, and A. Narayanan, "OpenWPM : An automated platform for web privacy measurement," 2015.

[32] M. Reed, P. Syverson, and D. Goldschlag, "Anonymous connections and onion routing," *IEEE J.*, 1998.

[33] A. Warren, "Tor Browser Artifacts in Windows 10," *SANS Inst. InfoSec Read. Room*, 2017.

[34] Google, "Google Public Policy Blog: Keep your opt-outs.," 2011. [Online]. Available: https://publicpolicy.googleblog.com/2011/01/keep-your-opt-outs.html.

[35] IETF, "Do Not Track: A Universal Third-Part," 2011.

[36] N. Kroes, "Why we need a sound Do-Not-Track standard for privacy online," 2012. [Online]. Available: https://www.politiekemonitor.nl/9353000/1/j9vvioaf0kku7zz/viwb6785rbw6.

[37] K.-W. Wu, S. Y. Huang, D. C. Yen, and I. Popova, "The effect of online privacy policy on consumer privacy concern and trust," *Comput. Human Behav.*, vol. 28, no. 3, pp. 889–897, May 2012.

[38] G. Kontaxis and A. D. Keromytis, "Protecting Insecure Communications with Topology-aware Network Tunnels," Department of Computer Science, Columbia University, USA, 2016.

[39] S. Englehardt and A. Narayanan, "Online Tracking: A 1-million-site Measurement and Analysis," *Proc. 2016 ACM SIGSAC Conf. Comput. Commun. Secur. - CCS'16*, no. 1, pp. 1388–1401, 2016.

[40] Joanna Geary, "DoubleClick: What is it and what does it do?," *The Guardian*, 2012. [Online]. Available: https://www.theguardian.com/technology/2012/apr/23/doubleclick-tracking-trackers-cookies-web-monitoring.

[41] Nicola Hughes, "AddThis (Clearspring): What is it and what does it do?," *The Guardian*, 2012. [Online]. Available: https://www.theguardian.com/technology/2012/apr/24/addthis-tracking-trackers-cookies-web-monitoring.

Anexos

A Lista de municípios visitados

Este anexo apresenta uma tabela com a lista dos 308 municípios visitados e o link de entrada no respetivo website.

Abrantes	http://www.cm-abrantes.pt/
Águeda	http://www.cm-agueda.pt/
Aguiar da Beira	http://www.cm-aguiardabeira.pt/
Alandroal	http://www.cm-alandroal.pt/pt/Paginas/home.aspx
Albergaria-a-Velha	http://www.cm-albergaria.pt/
Albufeira	http://www.cm-albufeira.pt/
Alcácer do Sal	http://www.cm-alcacerdosal.pt/pt/
Alcanena	http://www.cm-alcanena.pt/index.php/pt
Alcobaça	http://www.cm-alcobaca.pt/pt/Default.aspx
Alcochete	http://www.cm-alcochete.pt/
Alcoutim	http://www.cm-alcoutim.pt/pt/Default.aspx
Alenquer	http://www.cm-alenquer.pt/
Alfândega da Fé	http://www.cm-alfandegadafe.pt/
Alijó	http://www.cm-alijo.pt/
Aljezur	http://www.cm-aljezur.pt/pt/Default.aspx
Aljustrel	http://www.mun-aljustrel.pt/default.aspx
Almada	http://www.m-almada.pt/xportal/xmain?xpid=cmav2
Almeida	http://www.cm-almeida.pt/Paginas/default.aspx
Almeirim	http://www.cm-almeirim.pt/
Almodôvar	http://www.cm-almodovar.pt/
Alpiarça	http://www.cm-alpiarca.pt/

Alter do Chão	http://www.cm-alter-chao.pt/pt/
Alvaiázere	http://www.cm-alvaiazere.pt/
Alvito	http://www.cm-alvito.pt/pt/Default.aspx
Amadora	http://www.cm-amadora.pt/
Amarante	http://www.cm-amarante.pt/
Amares	http://www.cm-amares.pt/
Anadia	http://www.cm-anadia.pt/
Angra do Heroísmo	http://www.cmah.pt/
Ansião	http://www.cm-ansiao.pt/
Arcos de Valdevez	http://www.cmav.pt/
Arganil	http://www.cm-arganil.pt/
Armamar	http://www.cm-armamar.pt/
Arouca	http://www.cm-arouca.pt/portal/index.php
Arraiolos	http://www.cm-arraiolos.pt/pt/Paginas/home.aspx
Arronches	http://www.cm-arronches.pt/pt/
Arruda dos Vinhos	http://www.cm-arruda.pt/
Aveiro	http://www.cm-aveiro.pt/www/
Avis	http://www.cm-avis.pt/
Azambuja	http://www.cm-azambuja.pt/
Baião	http://www.cm-baiao.pt/
Barcelos	http://www.cm-barcelos.pt/
Barrancos	http://www.cm-barrancos.pt/
Barreiro	http://www.cm-barreiro.pt/
Batalha	http://www.cm-batalha.pt/
Beja	http://www.cm-beja.pt/homepage.do2
Belmonte	http://www.cm-belmonte.pt
Benavente	http://www.cm-benavente.pt/
Bombarral	http://www.cm-bombarral.pt/events/
Borba	http://www.cm-borba.pt/pt/Paginas/home.aspx
Boticas	http://www.cm-boticas.pt/

Braga	http://www.cm-braga.pt/pt
Bragança	http://www.cm-braganca.pt/
Cabeceiras de Basto	http://www.cabeceirasdebasto.pt/
Cadaval	http://www.cm-cadaval.pt/
Caldas da Rainha	http://www.cm-caldas-rainha.pt/
Calheta (Açores)	http://www.cmcalheta.pt/
Calheta (Madeira)	http://www.cm-calheta.pt/
Câmara de Lobos	http://www.cm-camaradelobos.pt/
Caminha	http://www.cm-caminha.pt/
Campo Maior	http://www.cm-campo-maior.pt/pt/
Cantanhede	http://www.cm-cantanhede.pt/mcsite/inicio/
Carrazeda de Ansiães	http://www.cm-carrazedadeansiaes.pt/
Carregal do Sal	http://www.carregal-digital.pt/
Cartaxo	http://www.cm-cartaxo.pt/Paginas/default.aspx
Cascais	http://www.cascais.pt/
Castanheira de Pera	http://www.cm-castanheiradepera.pt/
Castelo Branco	http://www.cm-castelobranco.pt/
Castelo de Paiva	http://www.cm-castelo-paiva.pt/
Castelo de Vide	http://www.cm-castelo-vide.pt/
Castro Daire	http://www.cm-castrodaire.pt/
Castro Marim	http://www.cm-castromarim.pt/site/
Castro Verde	http://www.cm-castroverde.pt/pt/Default.aspx
Celorico da Beira	http://www.cm-celoricodabeira.pt/Paginas/default.aspx
Celorico de Basto	http://www.mun-celoricodebasto.pt/
Chamusca	http://www.cm-chamusca.pt/
Chaves	http://www.chaves.pt/
Cinfães	http://www.cm-cinfaes.pt/
Coimbra	http://www.cm-coimbra.pt/
Condeixa-a-Nova	http://www.cm-condeixa.pt/
Constância	http://www.cm-constancia.pt/

Coruche	http://www.cm-coruche.pt/
Corvo	http://www.cm-corvo.pt/www/
Covilhã	http://www.cm-covilha.pt/
Crato	http://www.cm-crato.pt/pt/
Cuba	http://www.cm-cuba.pt/
Elvas	http://www.cm-elvas.pt/
Entroncamento	http://www.cm-entroncamento.pt/
Espinho	http://www.portal.cm-espinho.pt/pt/
Esposende	http://www.municipio.esposende.pt/
Estarreja	http://www.cm-estarreja.pt/
Estremoz	http://www.cm-estremoz.pt/
Évora	http://www.cm-evora.pt/pt/Paginas/home.aspx
Fafe	http://www.cm-fafe.pt/
Faro	http://www.cm-faro.pt/pt/Default.aspx
Felgueiras	http://www.cm-felgueiras.pt/
Ferreira do Alentejo	http://www.ferreiradoalentejo.pt/
Ferreira do Zêzere	http://www.cm-ferreiradozezere.pt/
Figueira da Foz	http://www.cm-figfoz.pt/
Figueira de Castelo Rodrigo	http://www.cm-fcr.pt/
Figueiró dos Vinhos	http://www.cm-figueirodosvinhos.pt
Fornos de Algodres	http://www.cm-fornosdealgodres.pt/
Freixo de Espada à Cinta	http://www.cm-freixoespadacinta.pt/
Fronteira	http://www.cm-fronteira.pt/pt/
Funchal	http://www.cm-funchal.pt/pt/
Fundão	http://www.cm-fundao.pt/
Gavião	http://www.cm-gaviao.pt/pt/
Góis	http://www.cm-gois.pt/
Golegã	http://www.cm-golega.pt/
Gondomar	http://www.cm-gondomar.pt/
Gouveia	http://www.cm-gouveia.pt/Paginas/default.aspx

Grândola	http://www.cm-grandola.pt/
Guarda	http://www.mun-guarda.pt/Portal/default.aspx
Guimarães	http://www.cm-guimaraes.pt/
Horta	http://www.cmhorta.pt/
Idanha-a-Nova	http://www.cm-idanhanova.pt/
Ílhavo	http://www.cm-ilhavo.pt/
Lagoa (Açores)	http://lagoa-acores.pt/
Lagoa (Algarve)	http://www.cm-lagoa.pt/index.php/pt/
Lagos	http://www.cm-lagos.pt/
Lajes das Flores	http://www.cmlajesdasflores.pt/
Lajes do Pico	http://www.cm-lajesdopico.pt/
Lamego	http://www.cm-lamego.pt/
Leiria	http://www.cm-leiria.pt/
Lisboa	http://www.cm-lisboa.pt/
Loulé	http://www.cm-loule.pt/
Loures	http://www.cm-loures.pt/
Lourinhã	http://www.cm-lourinha.pt/
Lousã	http://www.cm-lousa.pt/
Lousada	http://www.cm-lousada.pt/
Mação	http://www.cm-macao.pt/
Macedo de Cavaleiros	http://www.cm-macedodecavaleiros.pt/
Machico	http://www.cm-machico.pt/
Madalena	http://www.cm-madalena.pt/
Mafra	http://www.cm-mafra.pt/
Maia	http://www.cm-maia.pt/
Mangualde	http://www.cmmangualde.pt/
Manteigas	http://www.cm-manteigas.pt/
Marco de Canaveses	http://www.cm-marco-canaveses.pt/
Marinha Grande	http://www.cm-mgrande.pt/
Marvão	http://www.cm-marvao.pt/pt/

Matosinhos	http://www.cm-matosinhos.pt/
Mealhada	http://www.cm-mealhada.pt/
Mêda	http://www.cm-meda.pt/
Melgaço	http://www.cm-melgaco.pt/
Mértola	http://www.cm-mertola.pt/
Mesão Frio	http://www.cm-mesaofrio.pt/
Mira	http://www.cm-mira.pt/
Miranda do Corvo	http://www.cm-mirandadocorvo.pt/pt/Default.aspx
Miranda do Douro	http://www.cm-mdouro.pt/
Mirandela	http://www.cm-mirandela.pt/
Mogadouro	http://www.mogadouro.pt/
Moimenta da Beira	http://www.cm-moimenta.pt/
Moita	http://www.cm-moita.pt/
Monção	http://www.cm-moncao.pt
Monchique	http://www.cm-monchique.pt/pt/Default.aspx
Mondim de Basto	http://www.municipio.mondimdebasto.pt/
Monforte	http://www.cm-monforte.pt/index.php/pt/
Montalegre	http://www.cm-montalegre.pt/
Montemor-o-Novo	http://www.cm-montemornovo.pt/
Montemor-o-Velho	http://www.cm-montemorvelho.pt/
Montijo	http://www.mun-montijo.pt/
Mora	http://www.cm-mora.pt/pt/Paginas/home.aspx
Mortágua	http://www.cm-mortagua.pt/index.php
Moura	http://www.cm-moura.pt/
Mourão	http://www.cm-mourao.pt/pt/Paginas/home.aspx
Murça	http://www.cm-murca.pt/
Murtosa	http://www.cm-murtosa.pt/
Nazaré	http://www.cm-nazare.pt/pt
Nelas	http://www.cm-nelas.pt/
Nisa	http://www.cm-nisa.pt/

Nordeste	http://www.cmnordeste.pt/
Óbidos	http://www.cm-obidos.pt/
Odemira	http://www.cm-odemira.pt/
Odivelas	http://www.cm-odivelas.pt/
Oeiras	http://www.cm-oeiras.pt/pt/Paginas/default.aspx
Oleiros	http://www.cm-oleiros.pt/
Olhão	http://www.cm-olhao.pt/
Oliveira de Azeméis	http://www.cm-oaz.pt/
Oliveira de Frades	http://www.cm-ofrades.com/
Oliveira do Bairro	http://www.cm-olb.pt/PageGen.aspx
Oliveira do Hospital	http://www.cm-oliveiradohospital.pt/
Ourém	http://www.ourem.pt/
Ourique	http://www.cm-ourique.pt/pt/Default.aspx
Ovar	http://www.cm-ovar.pt/pt/Default.aspx
Paços de Ferreira	http://www.cm-pacosdeferreira.pt/
Palmela	http://www.cm-palmela.pt/
Pampilhosa da Serra	http://www.cm-pampilhosadaserra.pt/
Paredes	http://www.cm-paredes.pt/
Paredes de Coura	http://www.paredesdecoura.pt/
Pedrógão Grande	http://www.cm-pedrogaogrande.pt
Penacova	http://www.cm-penacova.pt/
Penafiel	http://www.cm-penafiel.pt/pt-pt/home.aspx
Penalva do Castelo	http://www.cm-penalvadocastelo.pt/
Penamacor	http://www.cm-penamacor.pt/cmp/
Penedono	http://www.cm-penedono.pt/
Penela	http://www.cm-penela.pt/
Peniche	http://www.cm-peniche.pt/
Peso da Régua	http://www.cm-pesoregua.pt/
Pinhel	http://www.cm-pinhel.pt/
Pombal	http://www.cm-pombal.pt/

Ponta Delgada	http://www.cm-pontadelgada.pt/
Ponta do Sol	http://www.cm-pontadosol.pt/
Ponte da Barca	http://www.cmpb.pt/ver.php?cod=0A0D0A
Ponte de Lima	http://www.cm-pontedelima.pt/
Ponte de Sor	http://www.cm-pontedesor.pt/
Portalegre	http://www.cm-portalegre.pt/pt/
Portel	http://www.cm-portel.pt/pt/Paginas/home.aspx
Portimão	http://www.cm-portimao.pt/
Porto	http://www.cm-porto.pt/
Porto de Mós	http://www.municipio-portodemos.pt/
Porto Moniz	http://www.portomoniz.pt/pt/
Porto Santo	http://www.cm-portosanto.pt/
Póvoa de Lanhoso	http://www.mun-planhoso.pt/
Póvoa de Varzim	http://www.cm-pvarzim.pt/
Povoação	http://www.cm-povoacao.pt/pvc/
Praia da Vitória	http://www.cmpv.pt/
Proença-a-Nova	http://www.cm-proencanova.pt/
Redondo	http://www.cm-redondo.pt/pt/paginas/home.aspx
Reguengos de Monsaraz	http://www.cm-reguengos-monsaraz.pt/
Resende	http://www.cm-resende.pt/
Ribeira Brava	http://www.cm-ribeirabrava.pt/cmrb1/
Ribeira de Pena	http://www.cm-rpena.pt/
Ribeira Grande	http://www.cm-ribeiragrande.pt/
Rio Maior	http://www.cm-riomaior.pt/
Sabrosa	http://www.sabrosa.pt/
Sabugal	http://www.cm-sabugal.pt/
Salvaterra de Magos	http://www.cm-salvaterrademagos.pt/
Santa Comba Dão	http://www.cm-santacombadao.pt/
Santa Cruz	http://www.cm-santacruz.pt/
Santa Cruz da Graciosa	http://www.cm-graciosa.pt/

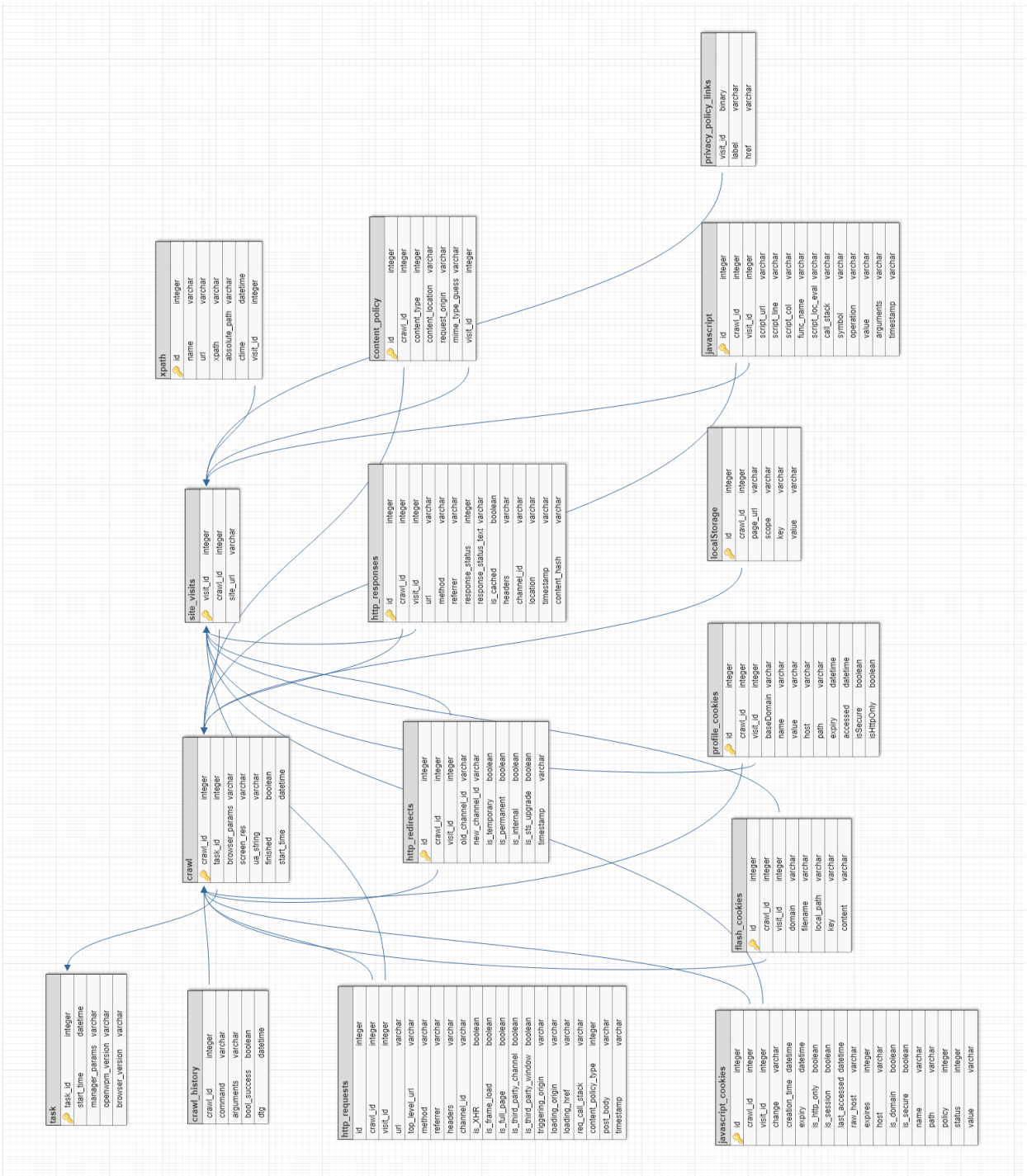
Santa Cruz das Flores	http://www.cmscflores.pt/
Santa Maria da Feira	http://www.cm-feira.pt/portal/site/cm-feira
Santa Marta de Penaguião	http://www.cm-smpenaguiao.pt/
Santana	http://www.cm-santana.com/pt/
Santarém	http://www.cm-santarem.pt/
Santiago do Cacém	http://www.cm-santiagocacem.pt/
Santo Tirso	http://www.cm-stirso.pt/
São Brás de Alportel	http://www.cm-sbras.pt/
São João da Madeira	http://www.cm-sjm.pt/
São João da Pesqueira	http://www.sjpesqueira.pt/
São Pedro do Sul	http://www.cm-spsul.pt/index.asp
São Roque do Pico	http://www.cm-saoroquedopico.pt/
São Vicente	http://www.cm-saovicente.pt/
Sardoal	http://www.cm-sardoal.pt/
Sátão	http://www.cm-satao.pt/
Seia	http://www.cm-seia.pt/
Seixal	http://www.cm-seixal.pt/
Sernancelhe	http://www.cm-sernancelhe.pt/
Serpa	http://www.cm-serpa.pt/
Sertã	http://www.cm-serta.pt/
Sesimbra	http://www.cm-sesimbra.pt/
Setúbal	http://www.mun-setubal.pt/
Sever do Vouga	http://www.cm-sever.pt/
Silves	http://www.cm-silves.pt/pt/Default.aspx
Sines	http://www.sines.pt/
Sintra	http://www.cm-sintra.pt/
Sobral de Monte Agraço	http://www.cm-sobral.pt/
Soure	http://www.cm-soure.pt/
Sousel	http://www.cm-sousel.pt/pt/
Tábua	http://www.cm-tabua.pt/

Tabuaço	http://www.cm-tabuaco.pt/
Tarouca	http://www.cm-tarouca.pt
Tavira	http://www.cm-tavira.pt/site/index.php
Terras de Bouro	http://www.cm-terrasdebouro.pt/
Tomar	http://www.cm-tomar.pt/
Tondela	http://www.cm-tondela.pt/
Torre de Moncorvo	http://www.cm-moncorvo.pt/
Torres Novas	http://www.cm-torresnovas.pt/
Torres Vedras	http://www.cm-tvedras.pt/
Trancoso	http://www.cm-trancoso.pt/
Trofa	http://www.mun-trofa.pt/
Vagos	http://www.cm-vagos.pt
Vale de Cambra	http://www.cm-valedecambra.pt/pages/1
Valença	http://www.cm-valenca.pt/
Valongo	http://www.cm-valongo.pt/
Valpaços	http://www.valpacos.pt/
Velas	http://www.cmvelas.pt/
Vendas Novas	http://www.cm-vendasnovas.pt/pt/Paginas/home.aspx
Viana do Alentejo	http://www.cm-vianadoalentejo.pt/
Viana do Castelo	http://www.cm-viana-castelo.pt/
Vidigueira	http://www.cm-vidigueira.pt/
Vieira do Minho	http://www.cm-vminho.pt/
Vila de Rei	http://www.cm-viladerei.pt/
Vila do Bispo	http://www.cm-viladobispo.pt/
Vila do Conde	http://www.cm-viladoconde.pt/
Vila do Porto	http://www.cm-viladoporto.pt/SITE/index.php
Vila Flor	http://www.cm-vilaflor.pt/
Vila Franca de Xira	http://www.cm-vfxira.pt/
Vila Franca do Campo	http://www.cmvfc.pt/
Vila Nova da Barquinha	http://www.cm-vnbarquinha.pt/

Vila Nova de Cerveira	http://www.cm-vncerveira.pt/
Vila Nova de Famalicão	http://www.cm-vnfamalicao.pt/
Vila Nova de Foz Côa	http://www.cm-fozcoa.pt/
Vila Nova de Gaia	http://www.cm-gaia.pt/pt/
Vila Nova de Paiva	http://www.cm-vnpaiva.pt/
Vila Nova de Poiares	http://www.cm-vilanovadepoiares.pt/
Vila Pouca de Aguiar	http://www.cm-vpaguiar.pt/
Vila Real	http://www.cm-vilareal.pt/
Vila Real de Santo António	http://www.cm-vrsa.pt/
Vila Velha de Ródão	http://www.cm-vvrodao.pt/
Vila Verde	http://www.cm-vilaverde.pt/
Vila Viçosa	http://www.cm-vilavicosa.pt/pt/Paginas/home.aspx
Vimioso	http://www.cm-vimioso.pt/
Vinhais	http://www.cm-vinhais.pt/
Viseu	http://www.cm-viseu.pt/
Vizela	http://www.cm-vizela.pt/
Vouzela	http://www.cm-vouzela.pt/

B Estrutura da Base de Dados

Este anexo ilustra a estrutura da base de dados SQLite onde estão registados os resultados obtidos.



C Lista de domínios de terceiros

Este anexo enumera as *cookies* que foram acedidas pelos diferentes domínios dos municípios.

- Addthis – as *cookies* deste domínio foram acedidas pelos seguintes municípios: Águeda, Alcobaça, Alcochete, Alcoutim, Alenquer, Alfandega da Fé, Aljezur, Aljustrel, Almada, Almodôvar, Alvito, Amadora, Ansião, Arcos de Valdevez, Aveiro, Barcelos, Barreiro, Boticas, Bragança, Câmara de Lobos, Carrazeda de Ansiães, Castelo Branco, Castro Verde, Chaves, Covilhã, Esposende, Fundão, Gondomar, Grândola, Guimarães, Idanha-a-Nova, Ílhavo, Leiria, Loulé, Macedo de Cavaleiros, Machico, Maia, Manteigas, Marinha Grande, Mealhada, Miranda do Corvo, Miranda do Douro, Mogadouro, Moimenta da Beira, Monchique, Montijo, Odemira, Oliveira de Azeméis, Ourique, Ovar, Palmela, Pampilhosa da Serra, Paredes, Ponta Delgada, Ponte da Barca, Ponte de Lima, Ródão, S. Brás de Alportel, Sabrosa, Santo Tirso, Sesimbra, Setúbal, Sever do Vouga, Sines, Tarouca, Terras de Bouro; Torres Vedras, Vagos, Valongo, Valpaços, Vila do Bispo, Vila do Conde, Vila Flor, Vila Franca de Xira, Vila Nova de Cerveira, Vila Real, Vila Real de Santo António, Vinhais e Vimioso;

- Farmácias de Serviços – as cookies deste domínio foram acedidas pelos seguintes municípios: Cadaval, Lourinhã, Amares, Ponte de Sôr, Elvas e Bombarral;

- Issuu – as cookies deste domínio foram acedidas pelos seguintes municípios: Vila Franca do Campo, Constância, Cuba, Velas e Vila Verde;

- Tempo – as cookies deste domínio foram acedidas pelos seguintes municípios: Portel, Lourinhã, Viana do Alentejo, Mora, Barrancos, Alvaiázere, Borba, Castro Verde, Vendas Novas, Celorico da Beira, Mourão, Nisa, Monforte e Arraiolos;

- Youtube – as cookies deste domínio foram acedidas pelos seguintes municípios: Carrazeda de Ansiães, Ribeira Brava, Cadaval, Câmara de Lobos, Sertã, Moura, Sabrosa, Figueira de Castelo Rodrigo, Batalha, Almeirim, Reguengos de Monsaraz, Belmonte, Ponte de Barca, Penacova, Amares, Seixal, Tavira, Amarante, Odemira, Gouveia, Monção, Cabeceiras de Basto, Beja, Arruda dos Vinhos, Redondo, Macedo de Cavaleiros, Lago, Pampilhosa da Serra, Oliveira do Hospital, Sesimbra, Vila Nova da Barquinha e S. Pedro do Sul;

- Zopim – as cookies deste domínio foram acedidas pelos seguintes municípios: Sever do Vouga, Carrazeda de Ansiães, Vimioso, Águeda, Arcos de Valdevez, Vila Nova de Cerveira, Albufeira, Odemira, Mirandela, Miranda do douro, Ponta Delgada, Alfândega da Fé, Cuba, Sabrosa, Moimenta da Beira, Valongo, Mogadouro, Vinhais e Esposende.